TBR WORKING PAPERS

Tennessee Board of Regents
Office of Policy & Strategy

The Impact of Undergraduate Research on Student Outcomes:

Examining High Impact Practices in TBR community colleges

Series on Student Engagement and High Impact Practices

January 2022

Iteration 2

Table of contents

Tables

Figures

The Impact of Undergraduate Research on Student Outcomes [1]

Abstract

We investigate whether student participation in undergraduate research—a high impact practice at TBR community colleges—has an effect on 1) academic performance; 2) probability of graduation, university transfer, and student departure; and 3) time to completion, transfer, and departure. Using enrollment, graduation, and course-taking data on community college students from the Tennessee Board of Regents and National Student Clearinghouse, we track the 2017 freshmen cohort over twelve calendar semesters and examine their exposure to undergraduate research experiences and their key educational outcomes. We generate propensity scores with a machine learning algorithm and use a doubly robust inverse probability weighting estimator to mitigate the selection bias and compare outcomes of similar students among undergraduate research participants and nonparticipants. Overall, we find that undergraduate research participants demonstrate better academic performance, are more likely to graduate and transfer, are less likely to depart, and progress to transfer and departure slower than similar nonparticipants. The effects grow in size with an increase in frequency of undergraduate research experiences.

Keywords: high impact practices, undergraduate research, community college, program evaluation, causal inference, propensity scores, generalized boosted modeling, dosage analysis, event history analysis

[1] The primary research team included **Alexander Gorbunov**, **Amy Moreland**, **Chris Tingle**, and **Russ Deaton**. The authors appreciate contribution and advice from Heidi Leming and Melynda Conner.

Introduction

This investigation is continuing a series of studies of student engagement and high impact practices (HIP) in the Tennessee Board of Regent (TBR) community colleges.[2] Undergraduate research is one of thirteen high impact practices that TBR has incrementally developed and expanded since 2015.[3]

Following the general framework developed for the study of service learning HIP (TBR, 2021), we examine the impact of participation in undergraduate research experiences (URE) on key college outcomes. Specifically, we investigate whether undergraduate research—as implemented at TBR community colleges—affects students' academic performance; the likelihood of completion, university transfer, and student departure; and progression to graduation, transfer, and departure. Participation in undergraduate research experiences and student outcomes are observed for a cohort of first-time freshmen, drawing on the data from the TBR data warehouse, the National Student Clearinghouse, and Tennessee Higher Education Commission. We examine both the general effect of undergraduate research participation and the impact of frequency of undergraduate research experiences. Although the TBR taxonomy differentiates between undergraduate research levels, we combine them in the overall undergraduate research experience to ensure large sample sizes in all models.

Descriptive analyses and statistical tests show that undergraduate research participants differ from nonparticipants on a host of important characteristics, creating a selection bias issue for the study. We mitigate the selection bias with an advanced approach, which we previously found to be superior to alternative methodologies in the study on the effect of service learning HIP on college outcomes (TBR, 2021). We find that participation in the undergraduate research HIP is significantly related to most outcomes of interest: final cumulative GPA, the probability of graduation, transfer, and departure; and time to transfer and departure. The effects differ by frequency of undergraduate research participation. We do not find any evidence that undergraduate research experiences affect time to graduation.

[2] The working papers and presentations on student engagement and high impact practices are available at https://www.tbr.edu/policy-strategy/presentations-and-papers
[3] TBR High Impact Practices are described at https://www.tbr.edu/student-success/tbr-high-impact-practices.

Literature review

This section briefly reviews a large body of extant scholarship on undergraduate research as an engaged learning experience or, as it is currently defined, a high impact practice. We start with key definitions, proceed to the literature on the university sector, and then take a closer look at undergraduate research implementation and studies of its effect on student outcomes in community colleges.[4]

Definitions

Undergraduate research experiences (URE) have been the focus of multiple studies. Although the exact interpretation of undergraduate research may differ depending on the institutional sector, time period, the discipline, professional organization, and the author, URE are generally defined as experiences that "use the scientific method and/or the engineering design process to promote student learning by investigating a problem where the solution is unknown to students or faculty" (Patton & Hause, 2020, p.3). The Association of American Colleges and Universities recognizes undergraduate research as a "high impact practice" (Kuh, 2008). High impact practices are defined as a set of evidence-based teaching and learning practices that engage students in deep learning over an extended period of time, are effective for student development and engagement, and are deemed (have been tested) to have a positive impact on student success in college (Kilgo et al, 2015; Kuh, 2008). The Tennessee Board of Regents (TBR), which recognizes and implements Undergraduate Research and Creative Activities as one of its high impact practices, applies the following *Minimum Definition of Practice*: "Undergraduate research is an inquiry or investigation conducted by an undergraduate student in collaboration with a faculty member that makes a unique intellectual, scholarly, or creative contribution to the discipline, and for which the student receives academic credit either through a course or independent study. The student's contribution may be part of a new or ongoing faculty research project (adapted from CUR)" (TBR, n.d.).[5]

---

[4] For a more comprehensive discussion of the origin and impact of high impact practices, please see the first TBR report on the topic (*The effect of service learning participation on college outcomes: An empirical investigation*), which is available here: https://www.tbr.edu/policy-strategy/presentations-and-papers

[5] TBR has developed detailed taxonomies of its high impact practices and introduced a system goal of having all students experience two HIPs before they complete a degree at a community college. As part of the HIP quality assurance, TBR has developed a "minimum definition of practice" for each HIP that it implements.

Undergraduate research in universities

The notion of research-based learning originated in the university sector; and undergraduate research, together with internships, has existed longer than any other engaged learning experience (Finley & McNair, 2013). As a pedagogy, undergraduate research came into lightlime when the Boyer Commission report (1998) came out. The renewed interest was partially driven by lack of growth in science education, stagnation in the number of STEM majors and degrees in science and engineering, recurring concerns about STEM workforce shortages, slipping US student rankings in science and mathematics and unfavorable international comparisons of shares of STEM graduates. As a result, the original conversations and various commission reports were focused around the STEM disciplines and science education in universities with the goal of STEM workforce development (Butz, 2004; Snyder et al., 2012; Lopatto, 2009; Olson, 2014). Consequently, in universities, undergraduate research was mostly implemented in science disciplines (Kuh, 2008).

Responding to the need to know the effect of undergraduate research on student outcomes, the number of studies of this pedagogical approach mushroomed. In one line of research, analysts conducted numerous case studies at individual institutions or a small number of campuses, often using ethnographic approaches or surveys, or mixed research methods to assess the effect of URE (Jonides et al., 1992; Ishiyama, 2001; Bauer & Bennett, 2003; Seymour et al. 2004). In the second line of research, large-scale surveys were developed and administered across multiple campuses, including Summer Undergraduate Research Experience (SURE, SURE II, SURE AY) and Classroom Undergraduate Research Experiences (CURE) [6] and others; some surveys were complemented by qualitative research (Lopatto, 2004, 2006, 2009; Seymour et al., 2004). Thanks to this body of scholarship, different effects of inquiry-based research on student outcomes have been studied in various contexts (Pascarella & Terenzini, 2005).

The watershed event happened in the second half of 2000's when two mutually reinforcing trends coalesced. Multiple studies examining the effect of undergraduate research participation on student

[6] The acronym CURE is also used in the field to denote course-based undergraduate research experience, which are discussed later in the literature review.

outcomes provided ample evidence for designating this experience as a high impact practice (Lopatto, 2009; Hewlett, 2021). In turn, official recognition of undergraduate research as a high impact practice by the Association of American Colleges and Universities (Kuh, 2008) created a new impetus for development of undergraduate research pedagogy and URE practices and proliferation of research on its implementation, student participation, and impacts on various student outcomes.

In the university sector, the following general outcomes were found to be related to URE: improved participation and persistence, better graduation outcomes, higher levels of student learning, gains in research skills, increased professional benefits, positive personal development changes, and an increased interest in, and higher likelihood of, enrolling in a graduate school (Nagda et al., 1998; Hathaway et al., 2002; Zydney et al., 2002; Bauer & Bennett, 2003; Seymour et al., 2004; Barlow & Villarejo, 2004; Lopatto, 2006, 2009; Kilgo et al., 2015; Verity et al., 2002; Brownell & Swaner, 2010; Eagan et al., 2013). A separate line of research examined early exposure to undergraduate research opportunities and found it to be beneficial to students (Stanford et al., 2017; Schuster, 2018; Wolkow et al., 2014; Rodenbusch et al., 2016; Nerio et al., 2019). Much attention has been paid to undergraduate research in sciences or specific STEM fields, with a focus on students and faculty (Nolan et al., 2020; Seymore et al., 2004; Lopatto, 2004; Jones et al., 2010; Zydney et al., 2002; Laursen et al., 2010). Positive effects of undergraduate research participation were documented for STEM students, especially for female, URM, and first-generation students (Balke et al., 2021). Participation in URE was found to be related to earning a STEM degree and graduation within six years for the STEM students (Rodenbusch et al., 2016).

A large body of scholarship focuses on the effects of undergraduate research participation for students from historically underrepresented groups, including minority, first-generation, and low income students (Jonides et al., 1992; Lopatto, 2004; Rodríguez et al., 2018; Summers & Hrabowski 2006; O'Donnell et al., 2015; Nagda et al., 1998; Ishiyama, 2001; Verity et al., 2002; Barlow & Villarejo, 2004; Thiry & Laursen, 2011; Eagan et al., 2013). Evaluating an undergraduate research program at the University of Michigan, Jonides et al., (1992) reported positive motivational and behavioral changes for underrepresented minority students. Nagda et al. (1998) found that undergraduate research participation

increases retention of minority students. Barlow and Villarejo (2004) reported increased persistence and graduation and a higher probability of enrolling in a graduate program for URE minority participants. In the online survey of 41 institutions, however, Lopatto (2004) did not find statistically significant differences between ethnic groups on reported gains and benefits and plans for graduate schools.

However, despite a large body of scholarship on the topic, there are certain caveats about generalizability of their findings and conclusions. As mentioned above, many of these investigations are case studies at individual institutions (liberal arts colleges or research universities with strong undergraduate research programs), which are often based on surveys of students and alumni, recruitment of participants, interviews, focus groups, and ethnographic observations (Zydney et al., 2002; Rodríguez et al., 2018; Bauer & Bennett, 2003; Thiry & Laursen, 2011; Nagda et al., 1998; Seymour, 2004; Lopatto, 2006; Hunter et al., 2007; Russell, 2005). The findings from this line of research may be limited to particular campuses, programs, and groups. Jonides et al. (1992) and Wolkow et al. (2014) used experimental setups to address their respective research questions; however, by design, their findings were limited to particular institutions. The above limitations present problems for generalizing the study findings to other student groups and other institutional and programmatic settings.

After reviewing sixty studies on URE, Linn et al. (2015) conclude that existing evidence for claims of positive effect is weak partly because much of the past research relies on self-reported gains in skills and learning collected via surveys and interviews. Eagan et al. (2013) concur: "many of these studies have serious shortcomings, which range from limited generalizability due to data collected from single-institution samples to over-estimation of the effect of undergraduate research programs by relying on simple descriptive statistics that fail to account for potential endogeneity in the data" (p.2). Hewlett (2016) echoes these concerns: "Despite the numerous studies touting the impact of the undergraduate research experience on students, there is currently a considerable debate within the educational community regarding the quality of the evidence and the ability of educational researchers to make legitimate causal claims about the impact of the UR experience on students" (p. 148).

A few quantitative or mixed-methods studies directly addressed the issue of selection bias. This issue stems from the systematic differences between students participating in URE and nonparticipants that introduce bias in the estimates. Using two national surveys, Eagan et al. (2013) employed propensity score matching in the propensity model and multinomial hierarchical generalized linear modeling in the output models to examine the relationship between URE participation and students' intentions to enroll in STEM graduate programs or non-STEM graduate programs relative to students with no plans for a graduate school. The researchers found that undergraduate research participation increases the probability of indicating plans to attend STEM-based graduate programs. Examining course-based undergraduate research experiences (CURE), Rodenbusch et al. (2016) also used propensity score matching to examine the impacts of early involvement of undergraduate students in research. They found that participants in a special CURE program at the University of Texas at Austin were more likely than nonparticipants to earn a STEM degree and graduate within 6 years; however, there was no evidence that participation affected final GPA. Nerio et al. (2019) used propensity score matching to examine outcomes of associate students, and their findings are discussed below in the review of the community college literature.

Undergraduate research in community colleges

Although community colleges began engaging their students in research at scale rather recently, implementation of, and collaboration around, undergraduate research experiences at community colleges have been developing fast. Researchers and analysts of these programs and initiatives are starting to follow suit (Cejda & Hensel, 2009; Hensel & Cejda, 2014; Brown et al., 2007; Schuster, 2017; Hewlett, 2016; 2018, 2021; CUR, 2021; Hensel, 2021).

Launched in 2005, the Community College Undergraduate Research Initiative (CCURI) currently incorporates multiple institutions in various states and helps colleges to integrate URE in their STEM programs (Hewlett, 2016; Bock & Hewlett, 2018).[7] The CCURI collaborates with the Council on Undergraduate Research (CUR) to provide assistance to two-year institutions in developing and

[7] The TBR's Volunteer State Community College is recognized as one of the most successful CCURI partners (Hewlett, 2016).

implementing undergraduate research experiences. The CCURI models include classroom-based URE as well as experiences with faculty acting as mentors for participating students (Nerio et al., 2019). However, for various reasons, the CCURI partners often cannot conduct rigorous educational research studies to assess the impact of URE participation on student outcomes (Hewlett, 2016).

In November 2019, the American Association of Community Colleges convened the Community College Undergraduate Research Experience Summit, which brought together 120 thought leaders with an aim of collecting their insight and planning for future expansion of URE opportunities in community colleges (Patton & Hause, 2020). The summit recognized the following examples of undergraduate research experiences that are offered by community colleges: course-based research, internships, independent studies, honors projects, STEM design challenges from real-word scenarios, competitions blending academic and technical skills, and mentored research, which is part of a larger project (p. 3). The body of scholarship on undergraduate research in two-year institutions is expanding. The Council on Undergraduate Research issued a volume on undergraduate research at community colleges (Hensel & Cejda, 2014) and, more recently devoted the entire issue of its journal *Scholarship and Practice of Undergraduate Research* to undergraduate research participation at community colleges (CUR, 2021).

Development of undergraduate research programs at community colleges presents a host of issues that stem, among other reasons, from the nature of the two-year institution. The traditional mission of the community college does not include research and production of scholarly knowledge, they are not deemed institutions in which either faculty or students engage in research, and they do not have graduate students who could serve as teaching assistants (Cejda & Hensel, 2009; Hensel & Cejda, 2014). Besides comparatively weaker research culture (Hewlett (2018) calls it “the lack of an undergraduate research culture altogether”), the other obstacles to URE implementation in community colleges include higher teaching expectations for faculty, fewer resources (including lack of space for some disciplines, infrastructure, institutional support, and funding for research), more diverse and less academically prepared student body than in research-intensive universities and small liberal arts colleges, and related logistical and competitive challenges (Hewlett, 2021; Hirst et al., 2014; Brown et al., 2007; Gentile et al.,

2017; Nolan et al., 2020). In addition, community colleges enroll greater shares of students from lower socioeconomic backgrounds and historically underrepresented groups, students with lack of academic preparedness, students who are enrolled part-time and are often transient, students who often work full-time and have family responsibilities, and students with limitations of financial aid and credit limits on their programs (Bangera & Brownell, 2014; Cejda & Hensel, 2009; Hensel & Cejda, 2014; Snyder & Cudney, 2017). Wolkow et al. (2014) summarize key issues as follows: “Perceived barriers of transformation at 2-yr institutions include heavier teaching responsibilities, resource and financial limitations, and higher representation of students who are at greater risk of failure” (p. 725).

Possible solutions to overcome some of these problems and expand opportunities for undergraduate research participation may include partnerships and collaboration with universities and intercampus connections among community colleges (Gentile et al., 2017; President’s Council of Advisors on Science and Technology, 2012; Hirst et al., 2014; Leggett-Robinson et al., 2015). Ashcroft et al. (2021), Gentile et al. (2017), and Russell et al. (2007) describe best practices in partnership between community colleges and universities to create impactful undergraduate research programs at community colleges. Course-based undergraduate research experiences (CUREs)—as opposed to independent research—are also suggested as an efficient way to enhance undergraduate research participation and diversity in community colleges especially if implemented at the introductory level (Gennet, 2021; Rodenbusch et al., 2016). CUREs are the courses that provide opportunities to all students to engage in research and earn credits rather than select few who apply for independent research experiences. However, there are few examples of successful CURE implementation (Bangera & Brownell, 2014; Wolkow et al., 2014).

Several studies examined the effect of URE participation on student outcomes in community colleges. The researchers discovered positive effect of undergraduate research on the following outcomes: higher learning gains, increased retention and graduation rates, higher GPA, greater participation in campus activities, and integration into the profession of the discipline (Foley & Leonhardt, 2017; Balke et al., 2021; Hewlett, 2016; Genet, 2021; Cejda & Hensel, 2009). However, similar to the above-mentioned caveats about the findings of studies in the university sector, most investigations of URE at community

colleges are cases studies of innovative practices and thus may have limited generalizability of findings to other institutional settings (Balke et al., 2021; Cejda & Hensel, 2009; Hensel & Cejda, 2014). Describing the research on URE conducted by the Community College Undergraduate Research Initiative, Hewlett (2016) notes that "much of what CCURI has been able to capture has relied on qualitative data generated from instruments such as structured interviews, surveys, and online data collection tools which include modified versions of the SUSSI (25) and SURE (26) surveys" (p. 148).

The largest study of the URE effect on associate student outcomes to date, which also addressed the selection bias issue via propensity score matching, was implemented by Nerio et al. (2019). These researchers employed mixed methodology to examine outcomes of STEM students who participated in the undergraduate research program in ten associate's degree granting colleges in the City University of New York (CUNY) system. They reported higher levels of retention in STEM disciplines, graduation with a STEM degree, and higher probability of transferring to research-intensive colleges and universities. However, the findings are not easily generalizable to other settings due to the idiosyncrasies of the community colleges in the CUNY system and the CUNY Research Scholars Program (CRSP) that was under investigation. For instance, the authors explain that at CUNY "community college faculty are governed by the same employment contract that serves faculty at 4-year schools and are expected to engage in scholarly activities to attain tenure and promotion despite having a greater teaching load. This has led to the adoption of a research culture… that is similar to the culture that exists in the system's eight 4-year schools and three comprehensive colleges" (p. 2). The CRPS program also has a very specific structure: a yearlong research program that offers faculty mentorship and a laboratory research project.

This study contributes to the existing research by conducting a longitudinal and systemwide examination of community college students across all disciplines, addressing the selection bias issue via an advanced technique, comparing key college outcomes of similar students among undergraduate research participants and nonparticipants, and examining the effect of both overall participation in URE and by frequency of undergraduate research experiences. Responding to the call by Lynn et al. (2015), we conduct a systematic study using multiple indicators of student success and appropriate method for each.

Dataset description

To address the research questions, this investigation linked several data sources: student enrollment, graduation, and HIP participation data from the TBR data warehouse; post-TBR enrollment and completion data from the National Student Clearinghouse (NSC); and data on the Pell Grant eligibility at any time and participation in the Tennessee Promise program from the Tennessee Student Assistance Corporation (TSAC).[8] The study has examined a cohort of first-time freshmen (including full- and part-time and summer-returning-fall students) enrolling in TBR community colleges in fall 2017 and tracked their undergraduate research participation and outcomes over time. The cohort was unduplicated by student ID: in case of simultaneous enrollment in two colleges in the first semester, the college with the highest number of attempted credits in that term was selected.

The study underwent two iterations. In Iteration 1, the observation period was limited to nine semesters, from fall 2017 through summer 2020.[9] In fall of 2021, the study was conducted again to include another year of the HIP participation and student tracking data (Iteration 2). Due to the data collection schedule, in fall 2021, the data on undergraduate research participation was available for eleven calendar terms: from fall 2017 through spring 2021. However, student outcomes—including the ones from both the TBR data warehouse and National Student Clearing House—were tracked from fall 2017 through the end of summer of 2021. Thus, the observation period for Iteration 2 of the study covers twelve calendar semesters or four full academic years.

The cohort includes 21,578 freshmen. Over the next eleven semesters, 3,300 students (15.3 percent) participated in undergraduate research. Among participants, 1,891 students (57.3 percent of participants) completed an undergraduate research component once, 753 students (22.8 percent) participated twice, and 656 students (19.9 percent) took part in this HIP three or more times. **Table 1** shows that most students (53.5 percent) participated in undergraduate research in their first and second semesters of enrollment.

[8] The authors express their gratitude to THEC/TSAC leadership for providing the Pell Grant eligibility and Tennessee Promise participation data.
[9] The previous version of the paper (Iteration 1 of the study) was posted on the TBR website with August 2021 publication month.

Table 1. Undergraduate research participation overall, by frequency, and over time

	Undergraduate Research *	Participation frequency categories **		
		Once	Twice	3+ times
	3,300	1,891	753	656
Fall 2017	22.3%	29.8%	27.1%	25.1%
Spring 2018	31.2%	24.0%	26.9%	27.2%
Summer 2018	1.8%	1.7%	1.4%	1.8%
Fall 2018	14.5%	14.9%	13.7%	12.5%
Spring 2019	16.2%	13.2%	14.8%	17.9%
Summer 2019	1.7%	1.7%	1.6%	1.5%
Fall 2019	5.1%	5.9%	5.9%	5.5%
Spring 2020	3.7%	3.9%	4.5%	4.9%
Summer 2020	0.5%	0.8%	0.5%	0.8%
Fall 2020	1.4%	2.1%	1.9%	1.5%
Spring 2021	1.7%	2.1%	1.8%	1.4%
	100%	100%	100%	100%

Notes. Percentage for each semester shows undergraduate research participation as a share of the overall participation over eleven terms. Duplication on ID is possible due to some students participating in multiple undergraduate research experiences over time.

* The data on undergraduate research participation was available through spring 2021.

** Frequency groups were determined based on the eventual classification of students over the entire observation period.

Table 2 presents demographic and academic variables for the cohort and by undergraduate research participation. It shows that students completing undergraduate research components differ from nonparticipants on several key characteristics. Statistically significant differences between participants and nonparticipants are observed across all racial/ethnic groups, except for the Other Race category. The difference is especially large for Black and White students: there are fewer Black students and more White students among participants (8.5 and 82 percent, respectively) than among nonparticipants (19.7 and 67.8 percent, respectively). The share of women in the cohort is higher both among participants (57.8 percent vs. 42.2 percent for men) and nonparticipants (55.8 percent as opposed to 44.2 percent for male students). Compared to nonparticipants, there are fewer adult (4.7 versus 8.1 percent) and Pell-eligible students (59.4 versus 65.4 percent) among students taking part in undergraduate research experiences.

Table 2 also shows that statistically significant differences between undergraduate research participants and nonparticipants are observed on all key academic variables. On average, HIP participants have higher high school GPA (3.22 versus 2.99), larger ACT composite score (20.1 versus 18.9), higher cumulative college GPA in the last semester of observation (2.66 versus 2.22), and more average cumulative credits earned (48.4 versus 31.3) than nonparticipants. Students engaging in undergraduate research tend to include fewer freshmen who need learning support than nonparticipants: 58.7 percent and 65.1 percent, respectively. Finally, the share of students who received assistance through the Tennessee Promise program is higher for undergraduate research participants (71 percent) than among those who did not take part in this HIP (61.7 percent).

Table 3 presents major groups at the time of the *first* undergraduate research experience. For some students, this major will be different from the one in their first term or during subsequent participation in this HIP. When students participate in undergraduate research, they tend to be enrolled in the following major fields: *Liberal Arts and Sciences, General Studies and Humanities* (77.6 percent); *Registered Nursing, Nursing Administration, Nursing Research and Clinical Nursing* (5.2 percent); *Education, General* (4.6 percent); *Business Administration, Management and Operations* (3.4 percent); *Computer and Information Sciences, General* (1.8 percent); and *Criminal Justice and Corrections* (1.2 percent).

Table 2. Participation in undergraduate research by demographic and academic variables

	Cohort	Undergraduate research participants		Nonparticipants	
		3,300		18,278	
		N	%	N	%
Asian #	314	30	0.91	284	1.55
Black #	3,872	281	8.52	3,591	19.65
Hispanic #	1,339	152	4.61	1,187	6.49
White #	15,093	2,705	81.97	12,388	67.78
Other	960	132	4.00	828	4.53
Male #	9,475	1,392	42.18	8,083	44.22
Female #	12,103	1,908	57.82	10,195	55.78
Traditional age #	19,941	3,145	95.30	16,796	91.89
Adult #	1,637	155	4.70	1,482	8.11
Non-Pell ever #	7,657	1,339	40.58	6,318	34.57
Pell ever #	13,921	1,961	59.42	11,960	65.43
Non-learning support #	7,749	1,362	41.27	6,387	34.94
Learning support #	13,829	1,938	58.73	11,891	65.06
Non-Promise #	7,952	958	29.03	6,994	38.26
Promise #	13,626	2,342	70.97	11,284	61.74
		Mean	Median	Mean	Median
High school GPA #		3.22	3.25	2.99	2.95
ACT score #		20.13	20.00	18.93	19.00
Final GPA #		2.66	2.8	2.22	2.04
Average credits earned #		48.39	57.00	31.33	23.50

Note. Percentage is of the total for each demographic breakdown.

# Indicates variables with a statistically significant difference between undergraduate research participants and nonparticipants (p<0.001) based on *chi-square* tests or two-independent samples *t*-tests for difference in means (two-sided).

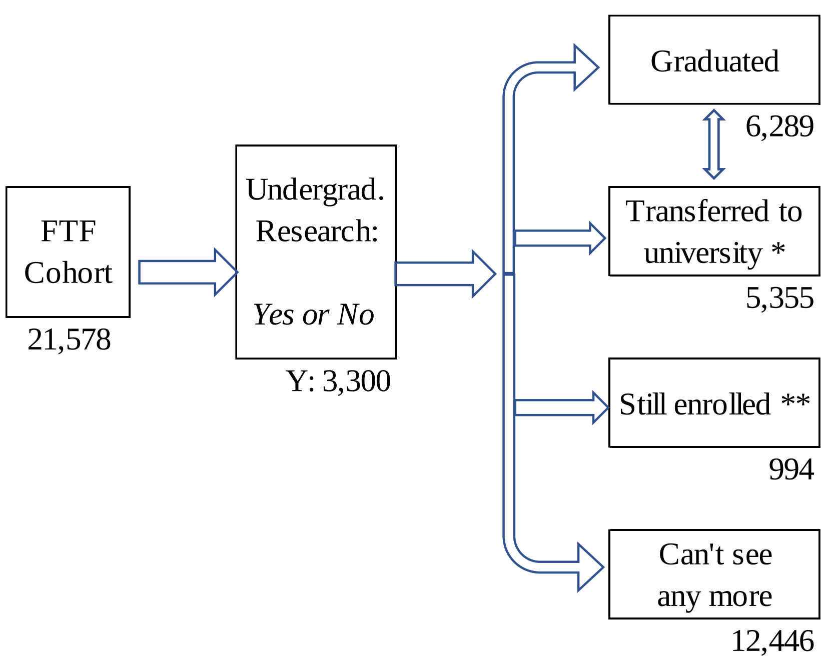
Table 3. Major in the first term of undergraduate research participation

Major	N	%
Liberal Arts and Sciences, General Studies and Humanities	3,854	77.61
Registered Nursing, Nursing Admin., Nursing Research & Clinical Nursing	259	5.22
Education, General	226	4.55
Business Administration, Management and Operations	171	3.44
Computer and Information Sciences, General	87	1.75
Criminal Justice and Corrections	59	1.19
Health Professions and Related Clinical Sciences, Other	57	1.15
Human Development, Family Studies, and Related Services	39	0.79
Allied Health Diagnostic, Intervention, and Treatment Professions	35	0.70
Industrial Production Technologies/Technicians	35	0.70
Electrical Engineering Technologies/Technicians	21	0.42
Design and Applied Arts	19	0.38
Audiovisual Communications Technologies/Technicians	11	0.22
Electromechanical Instrumentation and Maintenance Technologies/Technicians	11	0.22
Music	11	0.22
Hospitality Administration/Management	10	0.20
Business Operations Support and Assistant Services	8	0.16
Vehicle Maintenance and Repair Technologies	8	0.16
Health and Medical Administrative Services	7	0.14
Allied Health and Medical Assisting Services	5	0.10
Dental Support Services and Allied Professions	5	0.10
Legal Support Services	5	0.10
Engineering Technology, General	4	0.08
Heavy/Industrial Equipment Maintenance Technologies	4	0.08
Multi-/Interdisciplinary Studies, General	4	0.08
Fire Protection	3	0.06
Drafting/Design Engineering Technologies/Technicians	2	0.04
Engineering, General	2	0.04
Health Services/Allied Health/Health Sciences, General	2	0.04
Heating, Air Conditioning, Ventilation & Refrigeration Maintenance Technology	1	0.02
Medical Illustration and Informatics	1	0.02

Note. Duplication on ID is possible due to some students participating in multiple undergraduate research experiences over time.

The remainder of this section examines key outcomes for the entire cohort and by undergraduate research participation. **Figure 1** shows the main student outcomes for the 2017 first-time freshmen cohort tracked for twelve calendar semesters, from fall 2017 through summer 2021. Out of 21,578 freshmen in the cohort, 3,300 (15.3 percent) participated in the undergraduate research HIP. The outcomes are presented for the entire cohort regardless of undergraduate research experience. By summer 2021, 6,289 students (29.1 percent) earned a college credential and 5,355 students (24.8 percent) transferred to a four-year college or university.[10] Completion and transfer are not mutually exclusive outcomes, and 3,506 students (16.2 percent) both graduated and transferred to university; this dynamic is shown by the reverse arrow in the graph. Graduation, which is understood as earning a technical certificate or a degree, may precede or follow transfer. In spring or summer of 2021, 994 students (4.6 percent) were still enrolled in TBR community colleges.[11] Finally, 12,446 students (57.7 percent of the cohort) either dropped out or stopped out and were no longer available for observation.

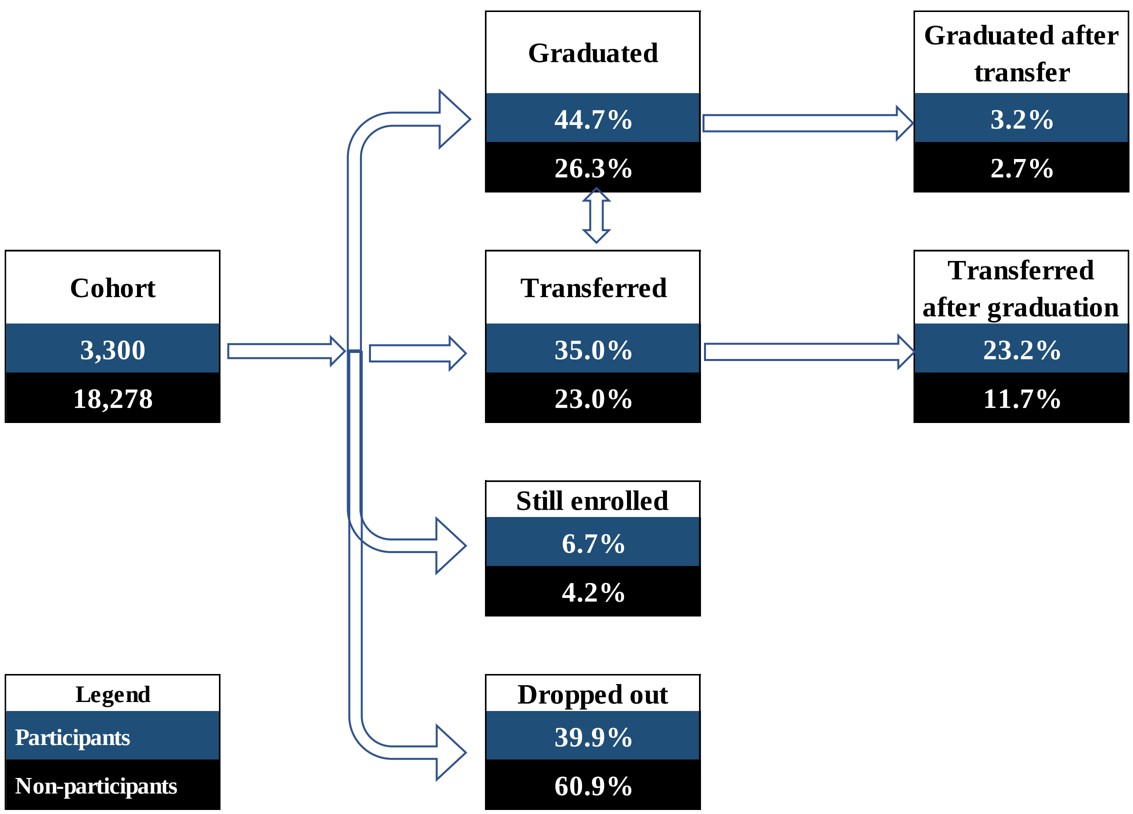
Figure 1. Outcomes for the 2017 first-time freshmen cohort



[10] Either in Tennessee or other states (limited to institutions submitting data to the National Student Clearinghouse).
[11] Iteration 1 used a more stringent definition of *Still Enrolled* students, which also used a minimum number of terms in the TBR system. Thus, the *Still Enrolled* and *Dropped Out* counts are not comparable across the iterations.

Figure 2 and **Table 4** present the main outcomes of the cohort by participation in undergraduate research experiences. They demonstrate that, at a purely descriptive level, undergraduate research participants and nonparticipants differ in their outcome attainment shares. As compared with students who do not take part in undergraduate research, larger shares of participants graduate (44.7 versus 26.3 percent), transfer to university (35 versus 23 percent), or are still enrolled at the end of the observation period (6.7 versus 4.2 percent). As a result, a smaller share of participants is among students who dropped out or stopped out: 39.9 percent as compared to 60.9 percent for nonparticipants. It is noteworthy that a much larger share of undergraduate research participants transfers to university after earning a community college credential: 23.2 percent of participants transfer after graduation as opposed to 11.7 percent of nonparticipants who do the same. This observation is related to the results of quantitative analyses examining progression to university transfer, which are presented in the *Results* section.

Figure 2. Outcomes by undergraduate research participation [12]



[12] Appendix 6 also breaks down descriptive outcomes by first-time URE participants in *Year 1* and *Year 2 or later*.

Table 4. Outcomes by undergraduate research participation

		Graduated		Transferred		Both outcomes	
	Total	All graduates	Did not transfer	All Transfers	Did not graduate	Graduation before transfer	Transfer before graduation
Participants	3,300	1,476	605	1,155	284	766	105
		44.73%	40.99%	35.00%	24.59%	23.21%	3.18%
Nonparticipants	18,278	4,813	2,178	4,200	1,565	2,146	489
		26.33%	45.25%	22.98%	37.26%	11.74%	2.68%

Note. Percentage is of the respective group's total.

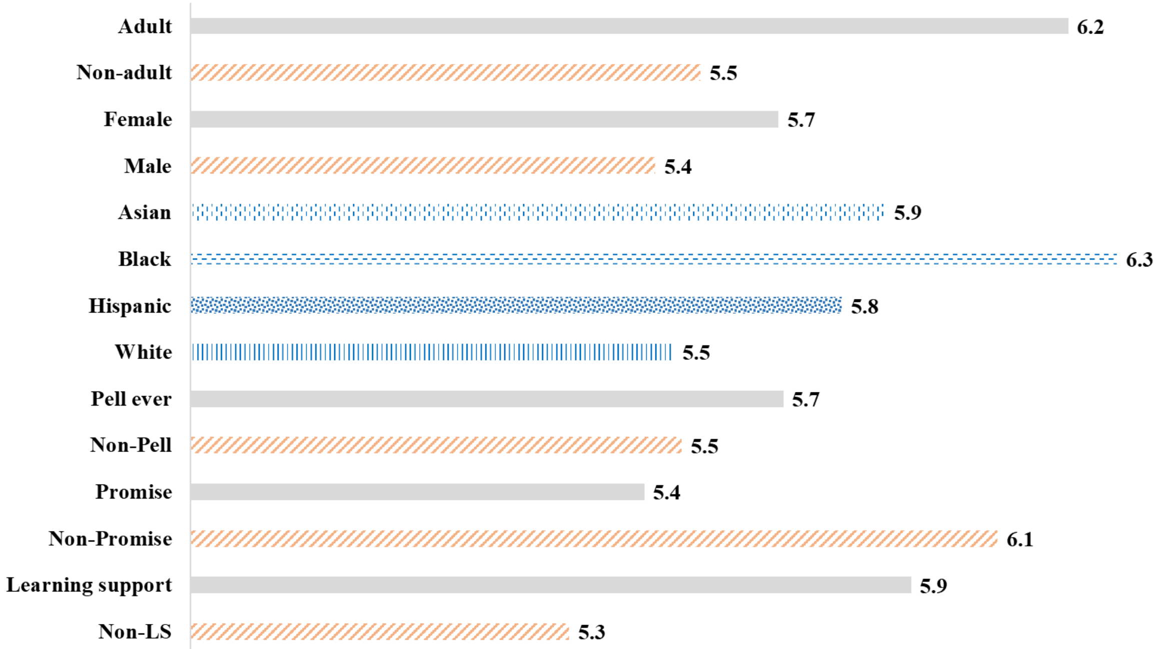
Table 5 presents the average time to graduation (any college credential versus associate degree or higher) and university transfer by undergraduate research participation, and by demographic and academic variables. Time to an outcome is measured both in semesters and cumulative attempted credits. For comparison, it also provides terms-to-outcome and credits-to-outcome for all graduates and transfer students in the cohort. **Figure 3** shows time to graduation in semesters by key demographic variables.

In semesters of enrollment, time to graduation with any award is similar for all comparison groups: on average, all graduates, undergraduate research participants, and nonparticipants take about 5.6 semesters to earn a college credential (either a technical certificate or a degree). However, participants attempt more credit hours before they graduate (72.7 credits as opposed to 70.5 for nonparticipants). In contrast, time to transfer shows more variability among the comparison groups: *nonparticipants* have the shortest average time (5.3 terms) and the smallest number of attempted credits (58.2 credit hours) to this outcome. Overall, undergraduate research participants have longer mean time to university transfer—measured in semesters (5.8 terms) or attempted credits (66.4 credit hours)—than nonparticipants or all transfer students. Similar to the above observation that undergraduate research participants transfer after graduation in much larger proportions than nonparticipants, this outcome is also related to the findings of the quantitative modeling of progression to university transfer in the *Results* section.

Table 5. Average time to graduation and university transfer in semesters and attempted credits

	Mean semesters and attempted credits to …					
	Graduation				Transfer	
	Any award (6,289)		Degrees (5,526)		(5,355)	
	Terms	Attempted credits	Terms	Attempted credits	Terms	Attempted credits
Total by outcome	5.57	71.0	5.71	71.3	5.40	60.0
UR participants	5.57	72.7	5.59	72.4	5.84	66.4
Nonparticipants	5.57	70.5	5.75	70.9	5.28	58.2
Demographic and academic variables						
Adult in Term 1	6.18	74.7	6.68	77.5	5.95	58.3
Trad. age in Term 1	5.52	70.9	5.65	71.0	5.38	60.0
Female	5.66	71.6	5.75	71.8	5.46	60.2
Male	5.44	70.2	5.65	70.5	5.33	59.6
Asian	5.85	72.4	5.91	71.2	5.97	64.8
Black	6.27	74.6	6.41	75.1	5.18	53.0
Hispanic	5.77	72.1	5.98	73.2	5.48	60.5
White	5.47	70.4	5.60	70.7	5.43	61.2
Other race	5.65	73.2	5.76	73.3	5.18	56.0
Pell ever	5.67	72.0	5.83	72.5	5.39	58.8
Non-Pell	5.48	70.1	5.60	70.2	5.41	60.9
Promise	5.42	71.3	5.53	71.4	5.54	63.1
Non-Promise	6.05	70.1	6.32	71.0	5.05	51.9
Learning support	5.90	72.1	6.12	73.0	5.46	57.8
Non-learning support	5.28	70.0	5.38	70.0	5.35	61.9

Figure 3. Time to graduation in semesters by demographic variables



In addition to time to outcome by undergraduate research participation and nonparticipation, Table 6 and Figure 3 also demonstrate how time to an outcome differs by demographic and academic variables. To illustrate, the following student groups have shorter average time to graduation than their counterparts: traditional-age, male and white students, individuals who were not Pell-eligible at any time, Tennessee Promise students, and students who did not require learning support during their first academic year of enrollment. Similar differences exist in credits to graduation (except for Promise students) and for university transfer as an outcome. Two exceptions are as follows: traditional-age students attempt slightly more credit hours by the time of their transfer than adult students and Black students have the shortest time to transfer measured either in semesters or attempted credits among all race/ethnicity groups.

These differences and the other results of the descriptive analysis demonstrate the need to account for demographic, academic, and other factors in quantitative models that aim to estimate the effect of undergraduate research participation on college outcomes. The *Methodology* section discusses the strategies that are used to address the research questions while accounting for these differences.

Methodology

This study investigates whether causal relationships exist between student participation in undergraduate research experiences and key college outcomes.[13] In the absence of random assignment to HIPs, direct comparison of outcomes of undergraduate research participants and nonparticipants will produce biased estimates due to selection bias. The selection bias is produced by the following factors: students can self-select into HIP participation (directly or due to their chosen major); as a result, participants and nonparticipants are expected to be systematically different on a number of observable and non-observable characteristics; and some of these differences will be related both to the treatment (undergraduate research participation) and the outcomes of interest. Under such conditions, the estimates from regular regression models comparing outcomes of participants and nonparticipants—even when controlling for a host of observable characteristics—will be biased.

We address the issue of selectivity by using the doubly robust augmented inverse propensity weighting (AIPW) method (Robins et al., 1994). In the subsequent *Results* section, the estimates from our final model using AIPW are compared with both the results of the naïve estimators, which do not handle selection into treatment, and estimates from the models using inverse probability of treatment weighting (IPTW) without adding the outcome model.

To estimate the propensity scores that are used in weighting, we apply the Generalized Boosted Regression (GBM) modeling. GBM uses a machine learning algorithm to estimate propensity scores. This algorithm fits multiple models using a boosted regression trees approach and then merges their predictions. GBM is deemed superior to alternative methods of estimating sample propensity scores and predicting treatment assignment According to Schonlau (2005), “There is a mounting empirical evidence that boosting is one of the best modeling approaches ever developed” (p. 331).

[13] This paper is the second quantitative study of the HIP effects on college outcomes. For a more comprehensive discussion of terminology, methodology, justification for using a machine learning approach for estimating propensity score (Generalized Boosted Modeling) and specific estimators (Event History Analysis), selection of covariates for propensity score models, alternative propensity score approaches, and handling data issues, please see the first TBR report (*The effect of service learning participation on college outcomes: An empirical investigation*), which examined the same cohort and is available here: https://www.tbr.edu/policy-strategy/presentations-and-papers

The GBM-estimated propensity scores are then used to reweight the observations via inverse probability of treatment weighting (IPTW). The IPTW approximates the mechanism of treatment assignment (HIP participation) as a function of the employed covariates. Imbens and Wooldridge (2009) showed that if the propensity scores are estimated correctly, using the IPTW is an effective method to neutralize the effect of selection on average treatment effect estimate. The main idea behind this method with propensity scores is to use probability weights to control for confounding. Using weights ensures that the distribution of confounders is the same for treated and untreated groups, which effectively results in removing these confounders. This approach weights more heavily untreated subjects who are more similar to treated cases and reduces the weights of untreated units who are dissimilar. The weight is estimated as the inverse of the propensity score (the probability of receiving the treatment). IPTW creates a pseudo-population that is representative of the treated group in terms of distribution of confounding factors. Stated differently, it constructs a comparison group of untreated subjects who are observationally similar to treated cases, and in the synthetic sample, the distribution of baseline covariates is independent of treatment assignment. The IPTW is used as the main estimator in this study due to its demonstrated ability to minimize bias, achieve balance on covariates, and yield statistically efficient estimates in comparison to other propensity score methods. In addition, IPTW is recommended for longitudinal data with time-varying confounders, which are present in this investigation (McCaffrey et al., 2004; Imbens, 2004; Austin, 2011; Austin & Stuart, 2015; Hirano et al., 2003; Thoemmes & Ong, 2015; Joffe et al., 2004; Imbens & Wooldridge, 2009; Sato & Matsuyama, 2003; Guo & Fraser, 2015).

As McCaffrey et al. (2004) state, “For a large sample size, the weighted treatment effect estimate will be nearly unbiased provided that several assumptions hold” (p. 405). To address the selection bias with IPTW, the following assumptions must be met: there are no unobserved confounding factors (observed covariates explain all pre-existing differences affecting outcomes); each subject has a nonzero probability (but not a probability of 1) to receive treatment; and the IPTW model is specified correctly (Thoemmes & Ong, 2015; McCaffrey et al., 2004). Although the first assumption is untestable, we select covariates based on prior literature, past research at TBR, and knowledge of factors that influence

participation in high impact practices. Thus, we argue that by modeling the relationship between observed covariates and treatment selection, we account for all key confounding factors. To make sure that IPTW helped correct for selection, we check balance on observed covariates using weighted standardized difference. Tabular and graphical assessment of balance in **Appendices 1**, **2**, and **3** demonstrate the effect of using IPTW on balance of observed covariates. The plausibility of the second assumption is assessed by examining if the empirical propensity score distributions overlap. We use boxplots of propensity scores to check for such an overlap between treated and untreated subjects in the propensity score space (**Appendix 4**). Finally, by using generalized boosted regression to estimate sample propensity scores, we address the issue of potential misspecification of the IPTW model.

Our final models use the augmented inverse propensity weighting (AIPW) method, which extends the IPTW method by including an outcome model. We employ a doubly robust estimator, which combines fitting models with inverse probability of treatment weights with the inclusion of additional pretreatment covariates (key demographic, academic, and other variables) to minimize any remaining bias. The doubly robust methods are consistent when either the outcome model or propensity model is correctly specified. Thus they provide another chance to correctly specify the model and—if either the propensity score model or the multivariate outcome model is specified correctly—to obtain more consistent and unbiased estimates of the treatment effect. So, it provides additional protection against misspecification of any model and gives two chances for a valid inference (Bang & Robins, 2005). Importantly, this methodology can be used—with respective modifications to account for nonbinary treatment—in the *dosage analysis*, and we use it to estimate effects of HIP participation frequency.

Finally, to avoid assigning too much weight to extreme cases, the weights are normalized and truncated at 1 percent and 99 percent of the weight distribution (Imbens, 2004; Cole & Hernan, 2008). To account for variability in the propensity score model and the fact that weights were estimated, we use robust (“sandwich”) standard errors to calculate the adjusted standard errors and confidence intervals for our estimates (Austin, 2011, 2016; Thoemmes & Ong, 2015). We use the following techniques to account for different outcomes and estimation issues: logistic and OLS regression and Event History Analysis.

Results

The *Methodology* section described the approach employed to estimate the effect of participation in undergraduate research experiences. Depending on the nature of the treatment variables, we used two primary strategies to obtain estimates that may have causal interpretation: 1) weighting on the inverse probability of treatment for binary analysis and 2) dosage analysis with generalized propensity scores for frequency effect examination. Both strategies were integrated with the "doubly robust" method to computing the estimates. Despite different weighting and modeling approaches, we present the estimates for both strategies together and discuss the findings for both types of analysis for each outcome. In other words, the outcomes of the binary and frequency analyses are provided together (i.e., in the same figure and table). The tables present the estimates for the unweighted and weighed samples and models with and without control variables. However, the subsequent discussion is mostly limited to statistically significant findings for the weighted samples from doubly robust models with control variables. The readers are advised to keep these explanations in mind when reading the interpretation of the results in this section.

The outcomes are presented in the following order. We start with OLS regression for final GPA and proceed to logistic regressions for the probability of graduation, transfer to university, and student departure. The following control variables were used in these models: age in the first semester, dummy variable for gender, race/ethnicity indicator variables, Pell-ever status flag, ACT composite score, and college in the first term. Then we present the outcomes of the Event History Analysis (EHA) approaches. These models estimate hazards for graduation, university transfer, and student departure and use the following control variables: age in each term, dummy variable for gender, race/ethnicity indicator variables, Pell-ever status flag, ACT composite score, and college of enrollment. Depending on the EHA model, outcome, and test results for proportional hazard assumption violation, some control variables were also modified as time-varying covariates (TVC) to include time-dependent effect (TVC are identified in the notes under the respective tables). All models estimate the average treatment effect

(ATE) using the inverse probability of treatment weights. Both for binary and frequency analysis, the normalized weights were truncated at 1 percent and 99 percent to attenuate the effect of large weights.

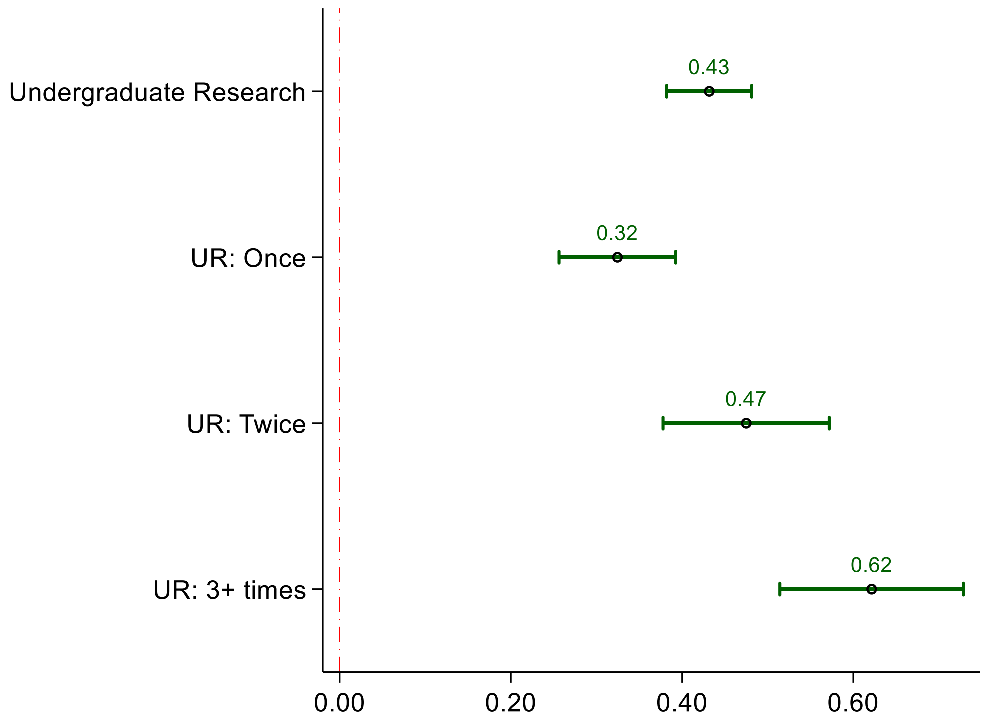
The first set of models estimates the effect of undergraduate research participation on final cumulative GPA (**Table 6** and **Figure 4**). To reiterate, we focus on the findings for the weighted samples from the doubly robust models with control variables. In all models, the coefficients are highly statistically significant and positive. We find that participation in undergraduate research is associated with a 0.43-point increase in final GPA, on average and holding everything else constant. The result for the weighted sample is attenuated when compared to the unweighted sample. This observation indicates the upward bias in the unweighted regression model, which does not address the selection bias issue. We also find that every frequency level of treatment—participating in undergraduate research once, twice, or three or more times during the observation period—is positively related to final GPA. Moreover, a rise in frequency leads to higher estimates of the treatment effect, ranging from 0.32-point increase for completing undergraduate research once to 0.62-point increase for participating three or more times.

Table 6. OLS estimates of impact of undergraduate research participation on final cumulative GPA

Treatment variables		Unweighted Sample		Weighted Sample #	
		Margin	SE	Margin	SE
Binary treatment (Reference group: nonparticipants)					
Undergraduate research		0.44 ***	(0.02)	0.34 ***	(0.02)
Undergraduate research (with control variables)		0.48 ***	(0.02)	0.43 ***	(0.03)
Non-binary treatment (Reference group: 0 times)					
Undergraduate research:	Once	0.40 ***	(0.02)	0.26 ***	(0.03)
Undergraduate research:	Twice	0.46 ***	(0.03)	0.34 ***	(0.04)
Undergraduate research:	Three + times	0.53 ***	(0.03)	0.42 ***	(0.04)
Undergraduate research:	Once (with controls)	0.43 ***	(0.02)	0.32 ***	(0.03)
Undergraduate research:	Twice (with controls)	0.54 ***	(0.03)	0.47 ***	(0.05)
Undergraduate research:	Three + times (with controls)	0.65 ***	(0.04)	0.62 ***	(0.05)

* p < .05, ** p < .01, *** p < .001. Robust standard errors in parentheses.
Control variables: age in the first term, gender, race/ethnicity groups, Pell-ever status, ACT composite score, college in first term.
Average Treatment Effect estimated with normalized Inverse Probability of Treatment Weights truncated at 1% and 99%.

Figure 4. Estimated effect of undergraduate research on final GPA: Binary and nonbinary treatments



The second set of models estimates the predicted probability of graduation due to undergraduate research participation (**Table 7** and **Figure 5**). As a reminder, graduation is interpreted as earning any postsecondary credential (technical certificate or degree) at TBR community colleges or other institutions after students left the TBR system. We find that participating in undergraduate research experiences increases the predicted probability of graduation by 21 percentage points. In frequency analysis, we find that undergraduate research with different frequency is positively related to the likelihood of earning a credential. The probability of graduation increases by 13 percentage points if students completed an undergraduate research component once, by 20 percentage points if they participated in it twice, and 34 percentage points if students were exposed to three or more undergraduate research experiences during the observation period. Again we see that the unweighted logistic regression has higher predicted probabilities in all models, indicating that weighting has minimized some positive bias in the estimates. Overall, we conclude that students who take part in undergraduate research experiences are more likely to graduate than their similar counterparts, and this effect grows with an increase in participation intensity.

Table 7. Predicted increase in the probability of graduation due to undergraduate research participation

Treatment variables	Unweighted Sample		Weighted Sample #	
	Margin	SE	Margin	SE
Binary treatment (Reference group: nonparticipants)				
Undergraduate research	0.82 ***	(0.04)	0.13 ***	(0.01)
Undergraduate research (with control variables)	1.30 ***	(0.06)	0.21 ***	(0.01)
Non-binary treatment (Reference group: 0 times)				
Undergraduate research: Once	0.16 ***	(0.01)	0.09 ***	(0.01)
Undergraduate research: Twice	0.17 ***	(0.02)	0.11 ***	(0.02)
Undergraduate research: Three + times	0.26 ***	(0.02)	0.21 ***	(0.02)
Undergraduate research: Once (with controls)	0.23 ***	(0.01)	0.13 ***	(0.02)
Undergraduate research: Twice (with controls)	0.29 ***	(0.02)	0.20 ***	(0.03)
Undergraduate research: Three + times (with controls)	0.40 ***	(0.02)	0.34 ***	(0.03)

* p < .05, ** p < .01, *** p < .001. Robust standard errors in parentheses.
Control variables: age in the first term, gender, race/ethnicity groups, Pell-ever status, ACT composite score, college in first term.
Average Treatment Effect estimated with normalized Inverse Probability of Treatment Weights truncated at 1% and 99%.

Figure 5. Predicted increase in probability of graduation: Binary and nonbinary treatments

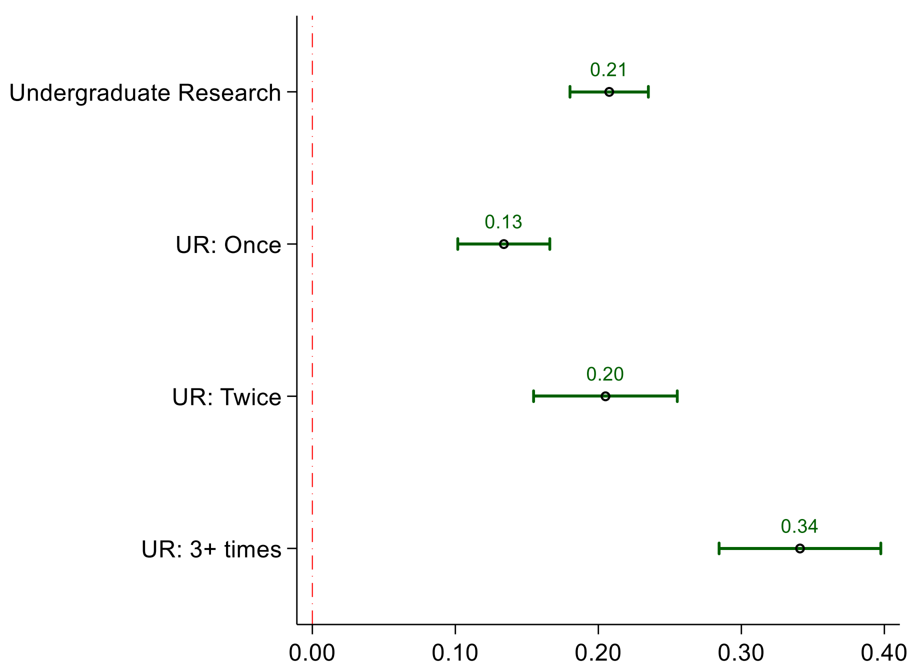
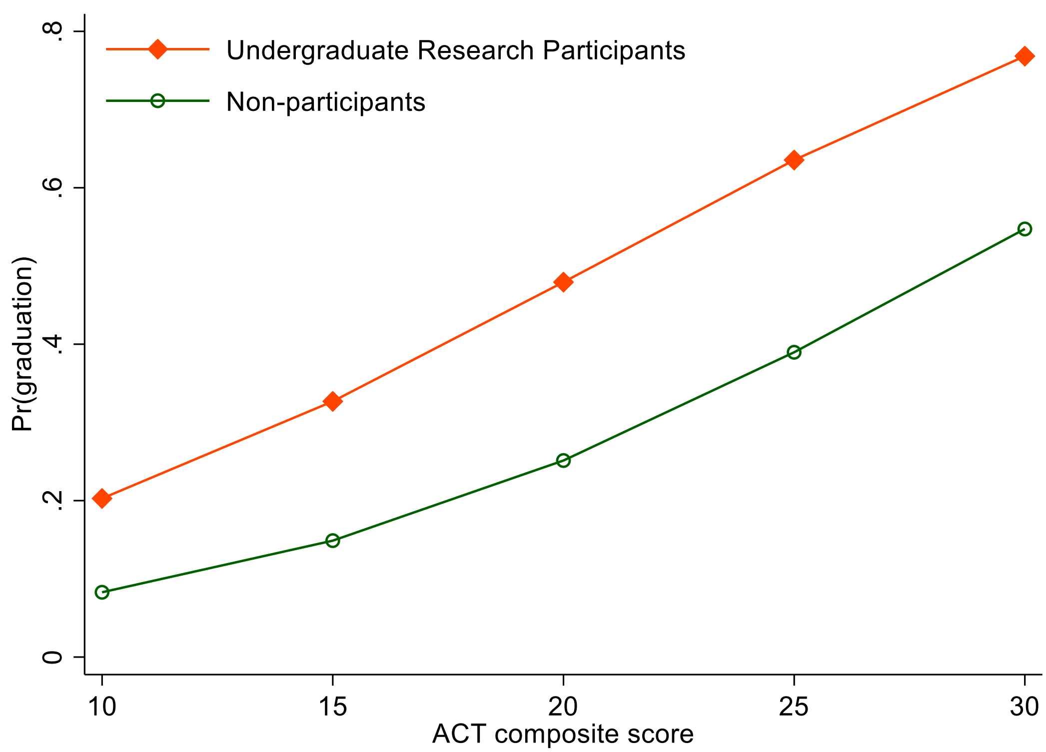


Figure 6 plots average marginal effects for undergraduate research participants and nonparticipants by ACT composite score in the logistic model for graduation with the binary treatment variable. Average marginal effects are predictions that are adjusted for actual observed values of covariates (not the mean values) in the model. In this approach to computing marginal effects, undergraduate research participants and nonparticipants differ only in the treatment exposure but are otherwise identical "average" subjects on all covariates. They are also similar on the inverse probability of treatment weights that are applied to the final model. Figure 6 shows that participants in undergraduate research experiences have a higher predicted probability of completion than nonparticipants for each ACT score, and the gap between these groups grows with an increase in the ACT composite score (except for the ACT score of 30 where the gap decreases slightly—likely due to a much smaller sample size).

Figure 6. Adjusted predictions for graduation for undergraduate research participants and nonparticipants



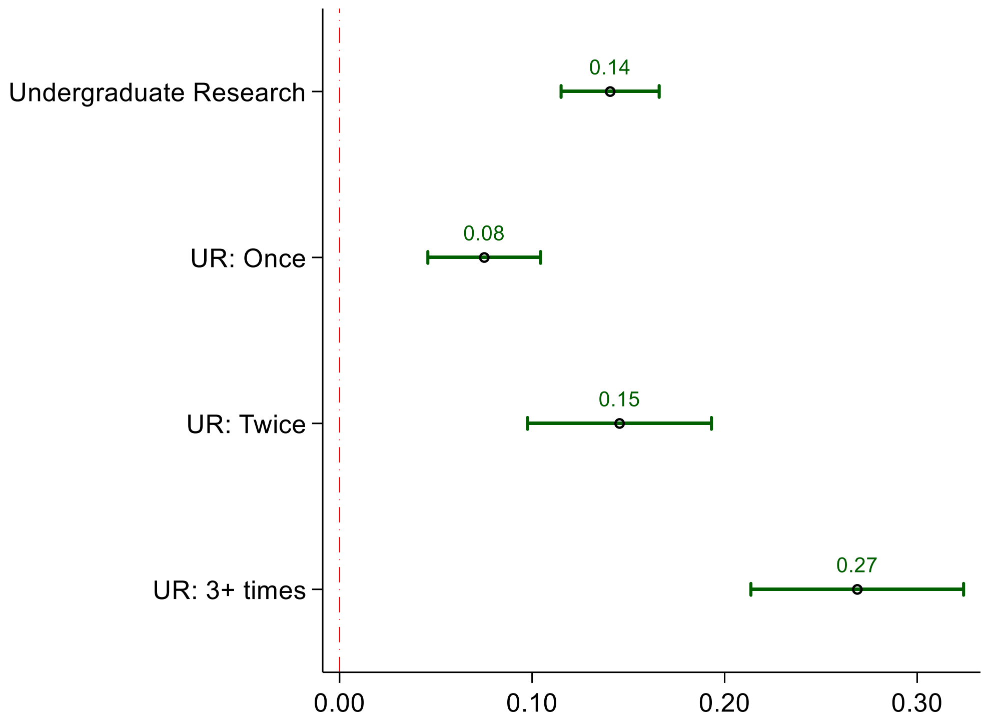
The third set of models examines the impact of undergraduate research participation on the probability of transfer to a four-year college or university. **Table 8** and **Figure 7** present all findings for this outcome for the unweighted and weighted samples, and binary and non-binary treatments. In all models, the coefficients are highly statistically significant and positive. Participation in undergraduate research is predicted to increase the probability of university transfer by 14 percentage points. We find that frequency of undergraduate research exposure has a positive effect on the likelihood of transfer, and the effect grows with each frequency level. Completing an undergraduate research component once increases the probability of transfer by 8 percentage points, participating in it twice by 15 percentage points, and taking advantage of three or more undergraduate research opportunities raises the predicted probability of transfer by 27 percentage points. Similar to the previous sets of models for GPA and graduation, the estimates in the weighted samples are smaller than the ones in the unweighted sample.

Table 8. Predicted increase in the probability of transfer due to undergraduate research participation

Treatment variables		Unweighted Sample		Weighted Sample #	
		Margin	SE	Margin	SE
Binary treatment (Reference group: nonparticipants)					
Undergraduate research		0.11 ***	(0.01)	0.08 ***	(0.01)
Undergraduate research (with control variables)		0.18 ***	(0.01)	0.14 ***	(0.01)
Non-binary treatment (Reference group: 0 times)					
Undergraduate research:	Once	0.09 ***	(0.01)	0.04***	(0.01)
Undergraduate research:	Twice	0.12 ***	(0.02)	0.08 ***	(0.02)
Undergraduate research:	Three + times	0.22 ***	(0.02)	0.18 ***	(0.02)
Undergraduate research:	Once (with controls)	0.16 ***	(0.01)	0.08 ***	(0.01)
Undergraduate research:	Twice (with controls)	0.25 ***	(0.02)	0.15 ***	(0.02)
Undergraduate research:	Three + times (with controls)	0.37 ***	(0.02)	0.27 ***	(0.03)

* p < .05, ** p < .01, *** p < .001. Robust standard errors in parentheses.
Control variables: age in the first term, gender, race/ethnicity groups, Pell-ever status, ACT composite score, college in first term.
Average Treatment Effect estimated with normalized Inverse Probability of Treatment Weights truncated at 1% and 99%.

Figure 7. Predicted increase in probability of transfer: Binary and nonbinary treatments



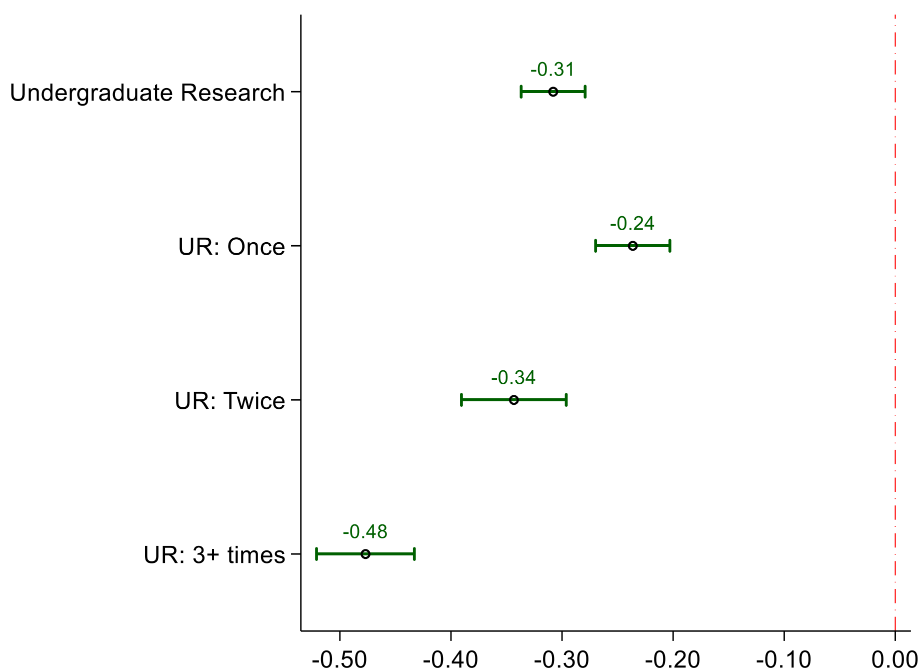
In our final set of logistic regression models, we examined the effect of undergraduate research participation on the probability of student departure. As a reminder, student departure is understood as dropping out or stopping out as of the end of the observation period. **Table 9** and **Figure 8** present the findings for this outcome for both binary and non-binary treatments. We find that undergraduate research participation has a strong negative statistically significant effect on the probability of student departure. Students completing undergraduate research components are 31 percentage points less likely to depart than their similar counterparts among nonparticipants. The dosage (frequency) analysis shows that the effect size grows with an increase in participation frequency. The predicted probability of departure decreases by 24 percentage points for community college students participating in undergraduate research once, 34 percentage points for students completing it twice, and 48 percentage points for students taking part in undergraduate research experiences three or more times as compared to similar nonparticipants.

Table 9. Predicted decrease in the probability of departure due to undergraduate research participation

Treatment variables		Unweighted Sample		Weighted Sample #	
		Margin	SE	Margin	SE
Binary treatment (Reference group: nonparticipants)					
Undergraduate research		- 0.85 ***	(0.04)	- 0.17 ***	(0.01)
Undergraduate research (with control variables)		- 1.44 ***	(0.06)	- 0.31 ***	(0.01)
Non-binary treatment (Reference group: 0 times)					
Undergraduate research:	Once	- 0.18 ***	(0.01)	- 0.12***	(0.01)
Undergraduate research:	Twice	- 0.20 ***	(0.02)	- 0.16 ***	(0.02)
Undergraduate research:	Three + times	- 0.29 ***	(0.02)	- 0.27 ***	(0.02)
Undergraduate research:	Once (with controls)	- 0.28 ***	(0.01)	- 0.24 ***	(0.02)
Undergraduate research:	Twice (with controls)	- 0.34 ***	(0.02)	- 0.34 ***	(0.02)
Undergraduate research:	Three + times (with controls)	- 0.44 ***	(0.02)	- 0.48 ***	(0.02)

* p < .05, ** p < .01, *** p < .001. Robust standard errors in parentheses.
Control variables: age in the first term, gender, race/ethnicity groups, Pell-ever status, ACT composite score, college in first term.
Average Treatment Effect estimated with normalized Inverse Probability of Treatment Weights truncated at 1% and 99%.

Figure 8. Predicted decrease in probability of departure: Binary and nonbinary treatments



The EHA models estimate hazards for three outcomes: graduation, university transfer, and student departure. Unlike the probability of an outcome in logistic models, hazards for graduation, transfer, and departure are estimated based on both whether the respective event took place and how long it took students to experience it. Thus, interpretation of outcomes in the logistic and EHA models is different. Similar to the models above, the EHA results are shown for both binary and non-binary treatments.

The first set of EHA models examines the effect of participation in undergraduate research on time to completion. **Table 10** and **Figure 9** provide estimates from the models for the hazard for graduation. Although the naïve models run on the unweighted sample produce statistically significant positive results for some treatment variables and frequency levels, our final model, which relies on the inverse probability of treatment weights and doubly robust estimators, does not find statistically significant results in any model specification. All confidence intervals in Figure 9 cross the line of zero effect, and we conclude that there is no evidence that undergraduate research participation accelerates progression to graduation.

Table 10. Cox proportional hazards model for time to graduation by undergraduate research participation

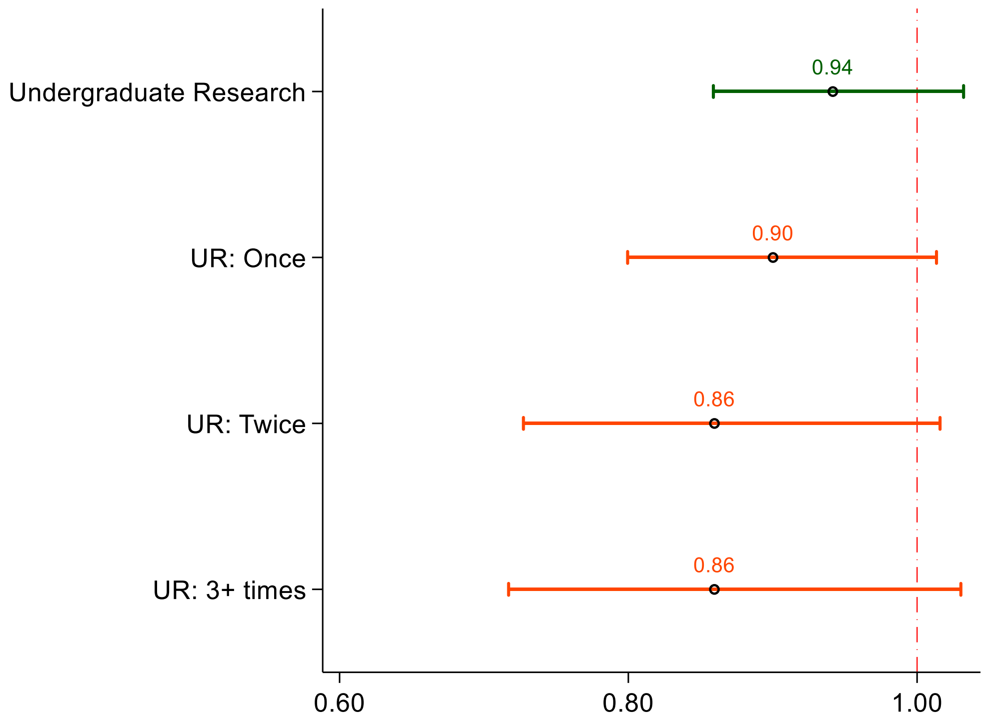
Treatment variables		Unweighted Sample		Weighted Sample #	
		Hazard ratio	95% CI	Hazard ratio	95% CI
Binary treatment (Reference group: nonparticipants)					
Undergraduate research		1.10 *	[1.04, 1.16]	0.93 *	[0.88, 0.99]
Undergraduate research (with control variables)		1.07	[0.99, 1.15]	0.94	[0.86, 1.03]
Non-binary treatment (Reference group: 0 times)					
Undergraduate research:	Once	1.10 *	[1.03, 1.18]	0.91 *	[0.84, 0.98]
Undergraduate research:	Twice	1.04	[0.94, 1.15]	0.88	[0.78, 1.01]
Undergraduate research:	Three + times	1.15 *	[1.04, 1.27]	0.98	[0.88, 1.09]
Undergraduate research:	Once (with controls)	1.10 *	[1.01, 1.20]	0.90	[0.80, 1.01]
Undergraduate research:	Twice (with controls)	1.00	[0.88, 1.13]	0.86	[0.73, 1.02]
Undergraduate research:	Three + times (with controls)	1.00	[0.87, 1.15]	0.86	[0.72, 1.03]

* Indicates statistically significant result (95% confidence interval does not include 1). Robust confidence intervals in brackets.

Control variables: age, gender, race/ethnicity groups, Pell status, ACT score, college of enrollment. Time-varying covariates in binary treatment models: age, gender, Black, Pell status, ACT score. Time-varying covariates in non-binary treatment models: gender, Asian, Black, White, Pell status, ACT score (based on tests of proportional hazard assumption violations).

Average Treatment Effect estimated with normalized Inverse Probability of Treatment Weights truncated at 1% and 99%.

Figure 9. Estimates of the effect on time to graduation: Binary and nonbinary treatments



The next set of EHA models estimates hazards for university transfer. **Table 11** shows hazard ratios from the respective model specifications, and **Figure 10** depicts the results graphically. Contrary to expectations, participation in undergraduate research is found to *decrease* hazard for university transfer. In the model with binary treatment variable, completing undergraduate research components decreases the hazard for transfer by 28 percent. (This decrease is the difference between 1 and the estimated hazard ratio of 0.72, which is 0.28). In frequency analysis models, the estimated effect of undergraduate research participation on the hazard for transfer grows with each frequency level. Namely, students taking part in undergraduate research once face the hazard for transfer that is 31 percent lower than the one for nonparticipants; students participating in this HIP twice have a 38 percent lower hazard for transfer, and the decrease in the transfer hazard for students experiencing undergraduate research three or more times is 40 percent. Overall, we conclude that undergraduate research participation slows down progression to university transfer. We offer a possible explanation for this finding in the *Discussion* section.

Table 11. Cox proportional hazards model for time to transfer by undergraduate research participation

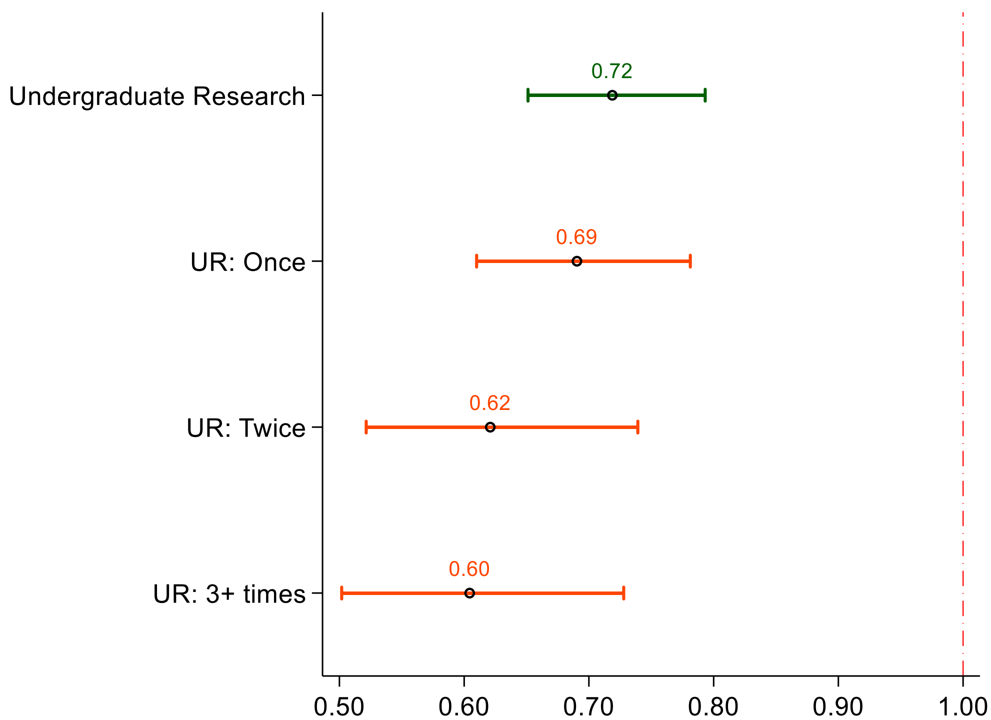
Treatment variables	Unweighted Sample		Weighted Sample #	
	Hazard ratio	95% CI	Hazard ratio	95% CI
Binary treatment (Reference group: nonparticipants)				
Undergraduate research	0.92 *	[0.87, 0.97]	0.79 *	[0.74, 0.84]
Undergraduate research (with control variables)	0.87 *	[0.79, 0.95]	0.72 *	[0.65, 0.79]
Non-binary treatment (Reference group: 0 times)				
Undergraduate research: Once	0.89 *	[0.82, 0.96]	0.77 *	[0.70, 0.84]
Undergraduate research: Twice	0.88 *	[0.79, 0.98]	0.75 *	[0.65, 0.86]
Undergraduate research: Three + times	1.03	[0.93, 1.13]	0.87 *	[0.78, 0.98]
Undergraduate research: Once (with controls)	0.88 *	[0.80, 0.97]	0.69 *	[0.61, 0.78]
Undergraduate research: Twice (with controls)	0.83 *	[0.72, 0.95]	0.62 *	[0.52, 0.74]
Undergraduate research: Three + times (with controls)	0.88	[0.75, 1.02]	0.60 *	[0.50, 0.73]

* Indicates statistically significant result (95% confidence interval does not include 1). Robust confidence intervals in brackets.

Control variables: age, gender, race/ethnicity groups, Pell status, ACT score, college of enrollment. Time-varying covariates in binary treatment models: age, Asian, Hispanic, White, Pell status. Time-varying covariates in non-binary treatment models: age, Asian, Black, White, ACT score (based on tests of proportional hazard assumption violations).

Average Treatment Effect estimated with normalized Inverse Probability of Treatment Weights truncated at 1% and 99%.

Figure 10. Estimates of the effect on time to transfer: Binary and nonbinary treatments



The final set of EHA models estimates the hazard for student departure for both binary and nonbinary treatment variables. **Table 12** and **Figure 11** provide hazard ratios for the respective model specifications. We find that undergraduate research participants are less likely to drop out or stop out in any semester than nonparticipants; the hazard for departure for the former decreases by 65 percent (1 – 0.35 = 0.65). In frequency analysis, participation in undergraduate research is also found to have a strong negative statistically significant effect on progression to student departure, and the effect size grows with each frequency level. Namely, students completing undergraduate research components one time face a hazard for departure that is 60 percent lower than that for nonparticipants. Students participating in two undergraduate research opportunities are 74 percent less likely to drop out or stop out in any semester than their counterparts. Finally, students who were exposed to three or more undergraduate research experiences during the observation period are very unlikely to drop out: the decrease in the departure hazard for them is 86 percent as compared to similar nonparticipants.

Table 12. Cox proportional hazards model for time to departure by undergraduate research participation

Treatment variables		Unweighted Sample		Weighted Sample #	
		Hazard ratio	95% CI	Hazard ratio	95% CI
Binary treatment (Reference group: nonparticipants)					
Undergraduate research		0.52 *	[0.50, 0.55]	0.57 *	[0.54, 0.60]
Undergraduate research (with control variables)		0.38 *	[0.36, 0.40]	0.35 *	[0.33, 0.38]
Non-binary treatment (Reference group: 0 times)					
Undergraduate research:	Once	0.58 *	[0.54, 0.62]	0.65 *	[0.61, 0.70]
Undergraduate research:	Twice	0.52 *	[0.47, 0.58]	0.56 *	[0.50, 0.63]
Undergraduate research:	Three + times	0.38 *	[0.34, 0.43]	0.39 *	[0.34, 0.45]
Undergraduate research:	Once (with controls)	0.44 *	[0.41, 0.47]	0.40 *	[0.37, 0.44]
Undergraduate research:	Twice (with controls)	0.35 *	[0.31, 0.38]	0.26 *	[0.22, 0.30]
Undergraduate research:	Three + times (with controls)	0.22 *	[0.19, 0.26]	0.14 *	[0.12, 0.17]

* Indicates statistically significant result (95% confidence interval does not include 1). Robust confidence intervals in brackets.

Control variables: age, gender, race/ethnicity groups, Pell status, ACT score, college of enrollment. Time-varying covariates in binary treatment models: age, gender, Black, Hispanic, Pell status. Time-varying covariates in non-binary treatment models: age, gender, Asian, Black, Hispanic, ACT score (based on tests of proportional hazard assumption violations).

Average Treatment Effect estimated with normalized Inverse Probability of Treatment Weights truncated at 1% and 99%.

Figure 11. Estimates of the effect on time to departure: Binary and nonbinary treatments

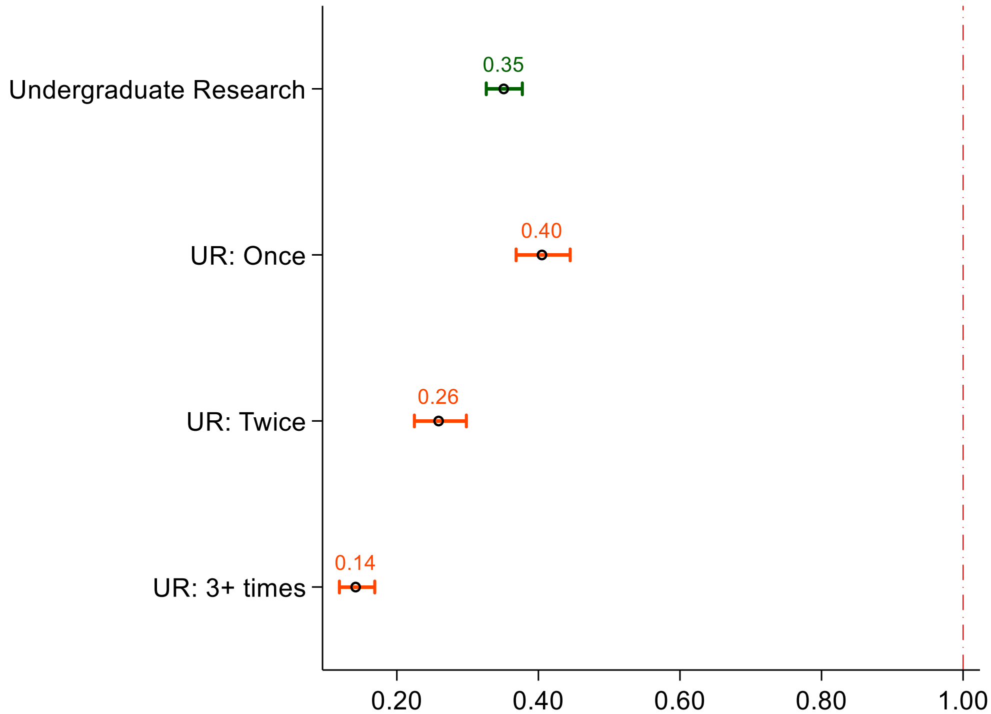


Table 13 summarizes all statistically significant findings of the study in a more interpretable way. Overall, we find that undergraduate research experiences (URE) are significantly related to most outcomes of interest: final cumulative GPA, the probability of graduation, university transfer, and student departure; and time to transfer and departure. The effects grow in size with an increase in frequency of undergraduate research participation. There is no evidence for the URE effect on time to completion.

Table 13. The impact of undergraduate research on student outcomes: Summary of findings

	Increase in GPA (points)	Change in predicted probability of:			Decrease in the hazard for:	
		Graduation	Transfer	Departure	Transfer	Departure
Overall UR	0.43	21 pp.	14 pp.	- 31 pp.	28%	65%
UR – once	0.32	13 pp.	8 pp.	- 24 pp.	31%	60%
UR – twice	0.47	20 pp.	15 pp.	- 34 pp.	38%	74%
UR – 3+ times	0.62	34 pp.	27 pp.	- 48 pp.	40%	86%

Discussion

This study contributes to prior literature on the impact of undergraduate research in several ways. First, we examined the effect of undergraduate research participation on multiple student outcomes in the community college settings and used appropriate methods for each outcome. Second, to address the self-selection bias, we employed an advanced bias-reducing method to obtain estimates that may have causal interpretation. The methodology uses the Generalized Boosted Regression modeling to estimate the probability of undergraduate research participation and doubly robust inverse probability weighting estimators to mitigate selection bias in our analysis of the HIP effect on student outcomes. Third, in addition to the general effect of completing an undergraduate research component, we investigated whether participation frequency affects the outcomes. Finally, we examined undergraduate research participation and student outcomes for a cohort of first-time freshmen across thirteen community colleges over twelve calendar semesters. Such a large-scale and systemwide approach, which also accounted for institutional effects and was not limited to specific majors, enabled us to overcome limitations of case studies of select institutions or disciplines, which may be affected by institutional or field idiosyncrasies.

We find that, as a high impact practice, undergraduate research experience positively influences several key student outcomes. As compared with similar nonparticipants, students completing undergraduate research components show better academic performance as measured by final GPA, are found to be more likely to graduate and transfer to a four-year college or university, are less likely to drop out or stop out, and progress slower to university transfer and departure. In frequency analysis, which examined the number of times this HIP was completed, we find that the effect size grows with an increase in frequency of undergraduate research experiences. Overall, we conclude that undergraduate research—as implemented by the TBR community colleges—is an efficacious high impact practice, which affects a multitude of educational outcomes. Our results are in line with prior literature (Nerio et al., 2019; Foley & Leonhardt, 2017; Balke et al., 2021; Hewlett, 2016; Genet, 2021; Cejda & Hensel, 2009) and provide an argument for expanding undergraduate research opportunities and increasing the experience frequency.

Our findings are mostly aligned with the original expectations and earlier literature. However, similar to the results of the study that examined the effects of service learning HIP,[14] one finding warrants additional attention. In contrast to the increased probability of university transfer as compared with similar nonparticipants, undergraduate research students also demonstrate longer time to transfer. In other words, we find that students who took part in the undergraduate research HIP progress to university transfer at a slower rate. It seems counterintuitive that increased probability of transfer in the logistic regression does not translate into faster progression to transfer in the Event History Analysis models. We offer the following explanation based on the descriptive analysis of the sample and survival function curves.[15] A larger share of undergraduate research participants transfer than nonparticipants (35 percent versus 23 percent); however, a larger proportion of participants transfer after earning a community college credential as compared to nonparticipants (23.2 percent versus 11.7 percent, respectively). Persisting in a community college until graduation and transferring after completion increases the time to university transfer for such students, including more participants. The average time to transfer is longer for participants than for nonparticipants (5.84 terms versus 5.28 terms, respectively); the former also attempt more credit hours by the transfer time than the latter (66.4 credits versus 58.2 credits). Thus, while undergraduate research participants are more likely to transfer than their similar counterparts, they also take longer to transfer to a four-year college or university—to a large extent because many of them attempt more credit hours and tend to earn a certificate or associate degree prior to transferring.

The study limitations are as follows. First, the conditional independence assumption (Angrist, 1997) postulates that treatment assignment (HIP participation) is independent of the outcome conditional on the predictors in the model. We argue that our approach and the treatment predictors used in the propensity and outcome models substantially mitigate the bias in the estimates. The predictor selection is rooted in prior research and our knowledge about the selection process into high impact practices at TBR. Based on that, we argue that our predictors adequately determine the selection mechanism and the outcomes. We

[14] TBR. (2021). *The effect of service learning participation on college outcomes: An empirical investigation*.
[15] Estimated survival curves for all outcomes in the weighted samples are presented in **Appendix 5**.

also check pre-weighting and post-weighting balances on these covariates and are satisfied with the results. Nevertheless, the conditional independence assumption is untestable, and our reliance on the observable pre-treatment characteristics is acknowledged as a study limitation.

Second, the study was conducted in community colleges of the Tennessee Board of Regents, which implements undergraduate research as a high impact practice in a centralized manner and alongside other HIPs and student engagement policies and practices. The TBR's work on expanding its HIPs and implementing other reforms has intensified over the past several years, and the observation period of the study captures these large-scale reform efforts. We also opted for a cohort-based approach to include only first-time freshmen students from a particular year. Therefore, our findings may not be easily generalizable to other state and institutional settings, which may use other approaches to undergraduate research implementation as well as other taxonomies and minimum definitions of practice.

Finally, the study is limited to a particular cohort, observation period, and a set of middle-range college outcomes. To preserve large sample sizes, we also used a combined identifier for undergraduate research participation without differentiating between types of student experiences in this HIP. When more data becomes available, it may be possible to extend the study horizon to more semesters of observation, compare outcomes of different cohorts and specific subpopulations of students, and examine distinct types of undergraduate research. Ideally, long-term educational and labor market outcomes should also be added to the investigation of HIP impacts. However, at the time of writing, the study is limited to the above parameters, and all results should be interpreted with these reservations in mind.

Next steps for the investigation of HIP effects stem from the study limitations. Future studies will use longer observation periods, rely on cohort comparisons, differentiate between undergraduate research experiences, examine mutual effects of other HIPs and TBR policies, and include additional long-term educational and labor market outcomes. Ideally, such quantitative analyses should be complemented by qualitative research and ethnographic studies, which may include interviews and focus groups of students, faculty, and alumni. Triangulating findings from different research approaches will provide for more in-depth understanding of the impact of participation in undergraduate research experiences and other HIPs.

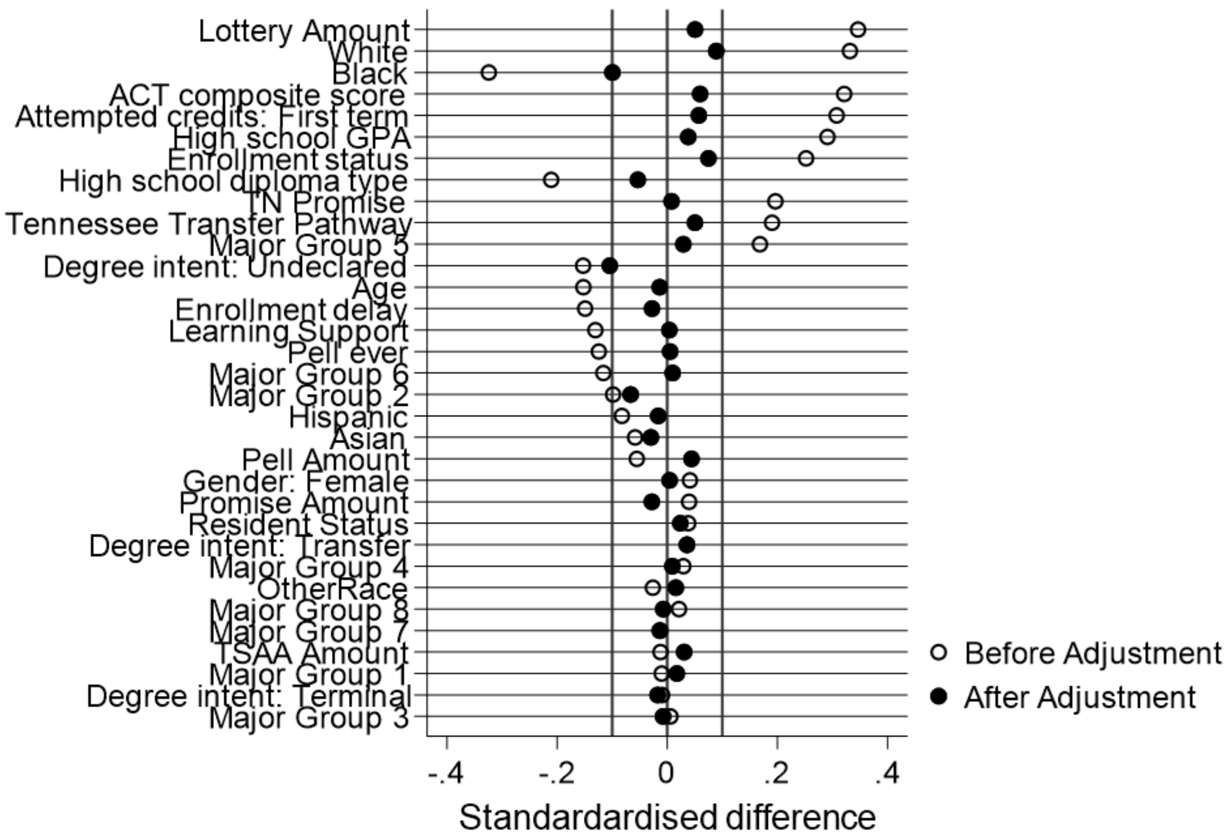
Appendices

Appendix 1. Standardized differences between mean values for treated and untreated groups before and after weighting, undergraduate research (binary treatment)

	Unweighted sample			Weighted sample		
	Participants	Non-participants	Stand. difference	Participants	Non-participants	Stand. difference
Age in the first semester	18.99	19.68	-0.152	19.52	19.58	-0.014
Gender: Female (0, 1)	0.58	0.56	0.041	0.56	0.56	0.004
Race/ethn.: Asian (0, 1)	0.01	0.02	-0.058	0.01	0.02	-0.03
Race/ethn.: Black (0, 1)	0.09	0.20	-0.324	0.15	0.18	-0.10
Race/ethn.: Hispanic (0, 1)	0.05	0.06	-0.083	0.06	0.06	-0.016
Race/ethn.: White (0, 1)	0.82	0.68	0.332	0.73	0.70	0.089
Race/ethn.: Other (0, 1))	0.04	0.05	-0.026	0.05	0.04	0.016
Learning Support (0, 1)	0.59	0.65	-0.131	0.64	0.64	0.004
ACT composite score	19.89	18.68	0.321	19.07	18.85	0.06
High school GPA	3.22	2.99	0.291	3.05	3.02	0.038
Resident Status	1.03	1.02	0.038	1.03	1.02	0.024
High school diploma type	1.76	1.92	-0.211	1.86	1.90	-0.053
Enrollment delay	0.85	1.45	-0.149	1.26	1.37	-0.028
Pell Amount	1,316.05	1,389.53	-0.055	1,440.29	1,381.93	0.044
Promise Amount	376.99	351.76	0.04	334.5	352.3	-0.028
Lottery Amount	895.61	627.35	0.346	704.11	664.97	0.051
TSAA Amount	217.92	221.67	-0.012	231.45	222.17	0.031
Pell-eligible ever (0, 1)	0.59	0.65	-0.124	0.65	0.65	0.005
Tennessee Promise (0, 1)	0.71	0.62	0.196	0.63	0.63	0.008
Degree intent: Transfer	0.66	0.64	0.036	0.66	0.64	0.036
Degree intent: Terminal	0.34	0.35	-0.009	0.34	0.34	-0.018
Degree intent: Undeclared	0.00	0.01	-0.153	0.00	0.01	-0.104
Enrollment status: Term 1	0.96	0.89	0.252	0.92	0.90	0.075
Attempted credits: Term 1	13.86	13.12	0.308	13.36	13.22	0.057
Transfer Pathway (0, 1)	0.27	0.19	0.19	0.22	0.20	0.05
Major Group 1 (0, 1)	0.00	0.01	-0.01	0.01	0.00	0.018
Major Group 2 (0, 1)	0.02	0.04	-0.098	0.02	0.03	-0.067
Major Group 3 (0, 1)	0.11	0.11	0.006	0.11	0.11	-0.008
Major Group 4 (0, 1)	0.04	0.03	0.029	0.03	0.03	0.009
Major Group 5 (0, 1)	0.21	0.15	0.168	0.17	0.16	0.029
Major Group 6 (0, 1)	0.44	0.50	-0.116	0.50	0.49	0.01
Major Group 7 (0, 1)	0.06	0.07	-0.013	0.06	0.07	-0.014
Major Group 8 (0, 1)	0.11	0.10	0.021	0.10	0.10	-0.007

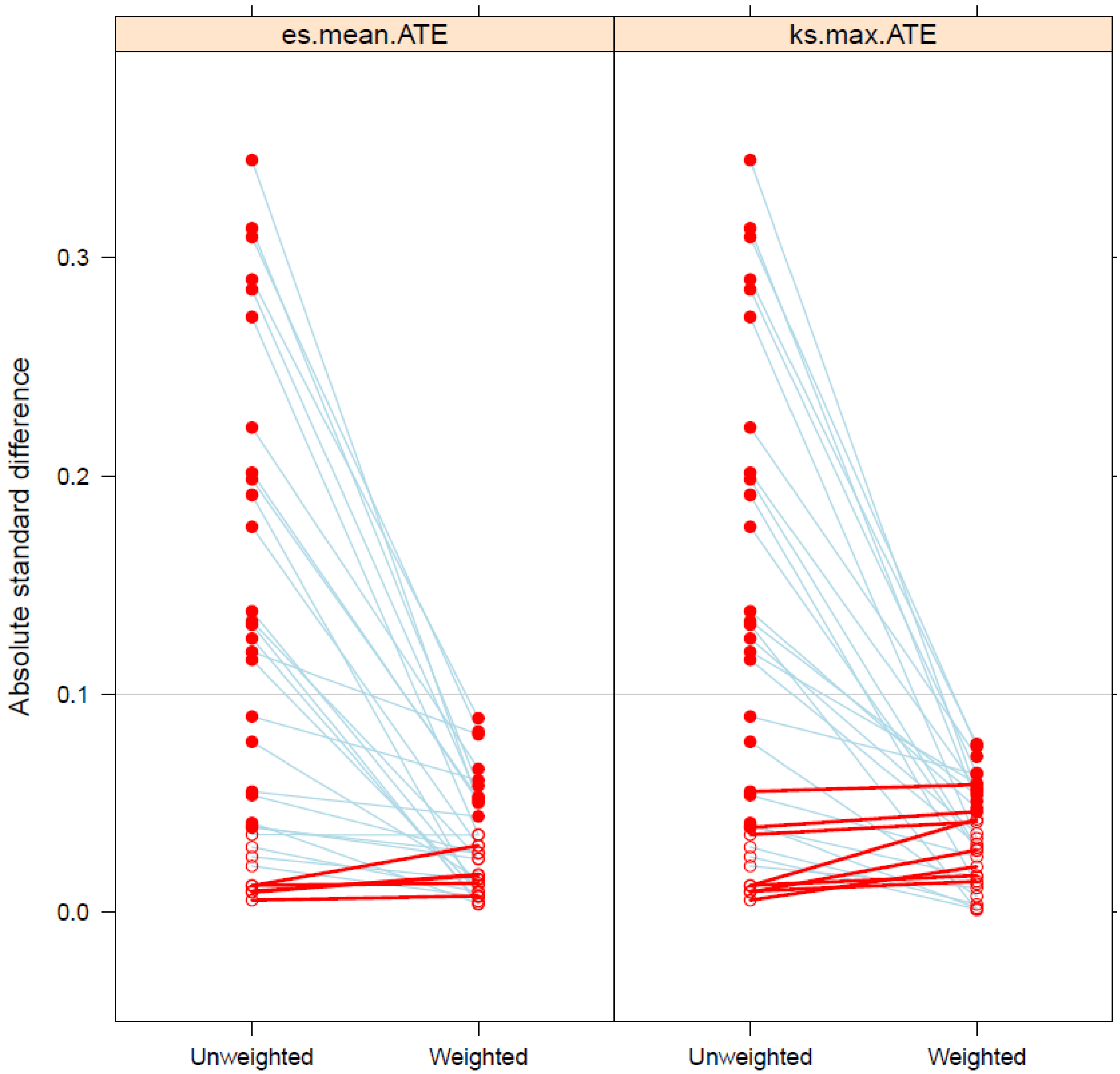
Note: An absolute standardized difference of ≤ 0.10 is considered balanced.

Appendix 2. Standardized difference between treated and untreated groups after ATE weighting, undergraduate research (binary treatment)



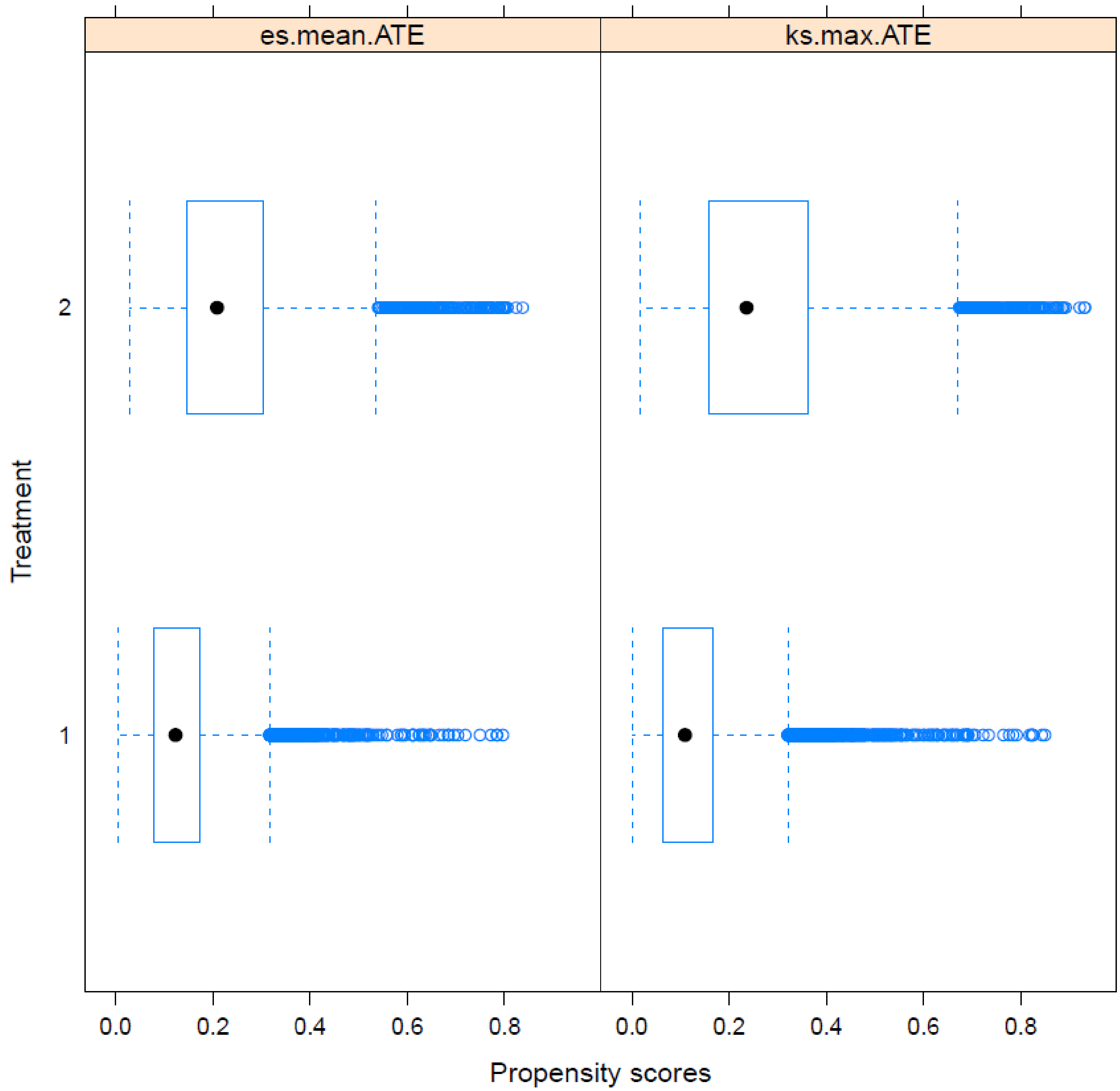
Note: An absolute standardized difference of ≤ 0.10 is considered balanced.

Appendix 3. Standardized difference between treated and untreated groups before and after ATE weighting, undergraduate research (binary treatment)



Note: A (absolute) standardized difference of ≤ 0.10 is considered balanced.

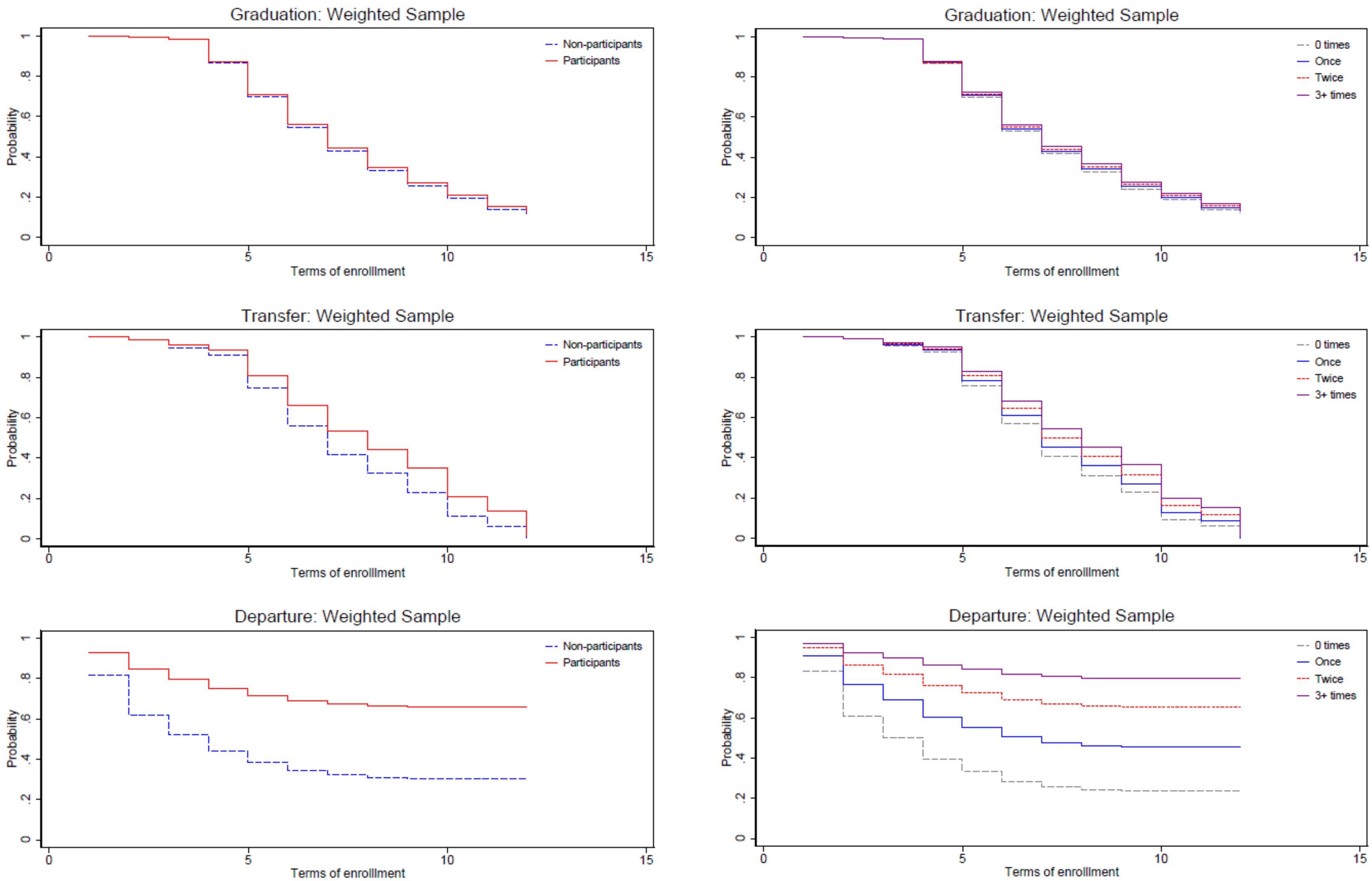
Appendix 4. Boxplot of propensity scores



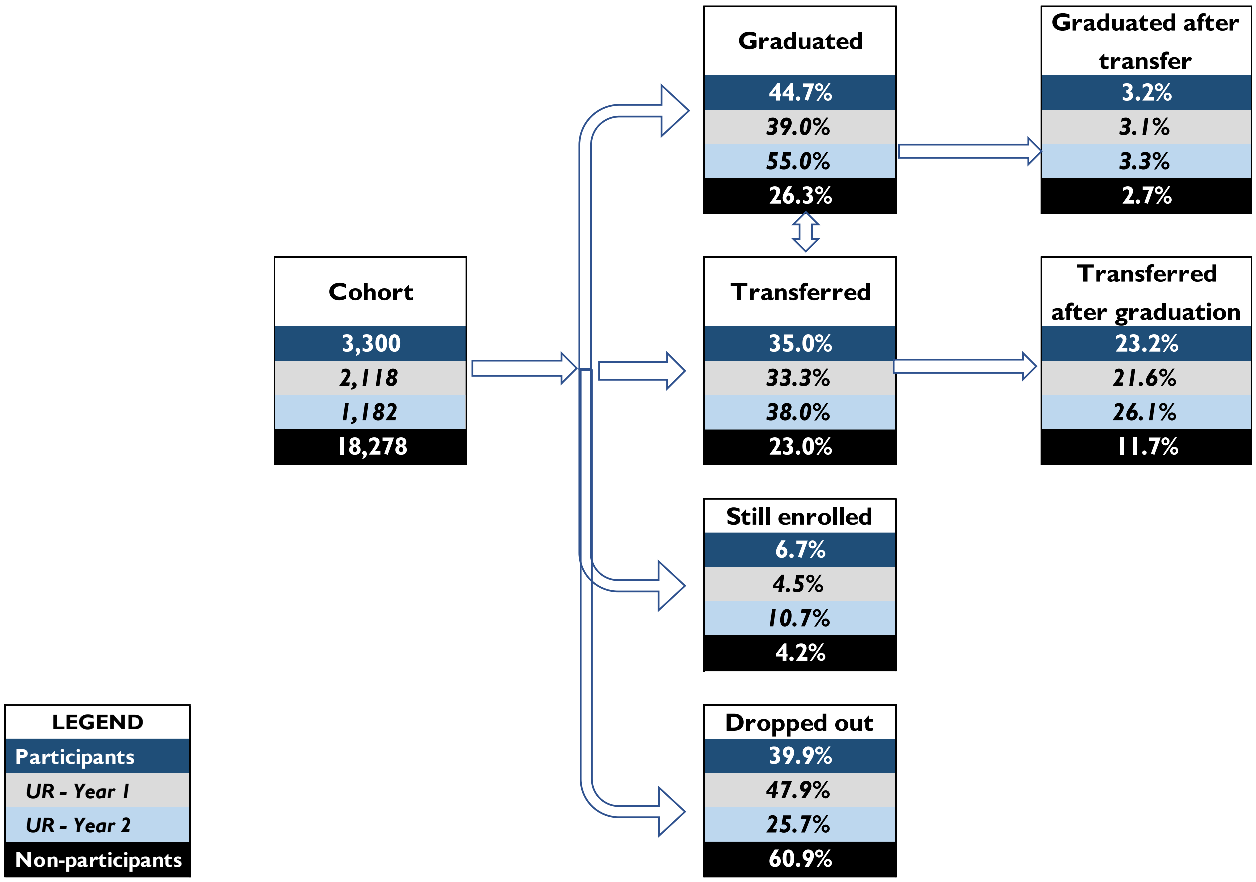
Appendix 5. Estimated survival curves by undergraduate research participation: Weighed samples

BINARY TREATMENT

NON-BINARY TREATMENT



Appendix 6. Descriptive outcomes overall and by the year of first undergraduate research participation



Note: Descriptive outcomes for UR participants in *Year 1* and *Year 2 or later* should be compared with caution as the latter are closer to attaining outcomes

References

Angrist, J. D. (1997). Conditional independence in sample selection models. *Economics Letters*, *54*(2), 103–112. https://doi.org/10.1016/S0165-1765(97)00022-0

Ashcroft, J., Jaramillo, V., Blatti, J., Guan, S.-S. A., Bui, A., Villasenor, V., Adamian, A., Chavira, G., Saetermoe, C., le, Eileen; Horowitz, B., Palacios, B., Williams, P. B., Brockman, E., & Moses, J. (2021). BUILDing equity in STEM: A collaborative undergraduate research program to increase achievement of underserved community college students. *Scholarship and Practice of Undergraduate Research*, *4*(3), 47-58. Council on Undergraduate Research.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399-424. https://doi.org/10.1080/00273171.2011.568786

Austin, P. C. (2016). Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. *Statistics in Medicine*, *35*(3), 5642–55. https://doi.org/10.1002/sim.7084

Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, *34*, 3661-3679. https://doi.org/10.1002/sim.6607

Balke, V., Grusenmeyer, L., & McDowel, J. (2021). Long-term outcomes of biotechnology student participation in undergraduate research experiences at Delaware Technical Community College. *Scholarship and Practice of Undergraduate Research*, *4*(3), 5-12. Council on Undergraduate Research. doi: 10.18833/spur/4/3/12

Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, *61*(4), 962–973. https://doi.org/10.1111/j.1541-0420.2005.00377.x

Bangera, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE–Life Sciences Education* 13, 602–606. doi: 10.1187/cbe.14-06-0099

Barlow A. E. L., & Villarejo M. (2004). Making a difference for minorities: Evaluation of an educational enrichment program. *Journal of Research in Science Teaching*, *41*(9), 861–881. https://doi.org/10.1002/tea.20029

Bauer, K., & Bennett, J. (2003). Alumni perceptions used to assess undergraduate research experience. *Journal of Higher Education*, *74*(2), 210-230. https://doi.org/10.1353/jhe.2003.0011

Bock, H., & Hewlett, J. (2018). Undergraduate research: Changing the culture at community colleges. *The Scholarship and Practice of Undergraduate Research*, *1*(4), 49–50.

Boyer Commission on Educating Undergraduates in the Research University. (1998). *Reinventing undergraduate education: A blueprint for America's research universities*. https://files.eric.ed.gov/fulltext/ED424840.pdf

Brown, D. R., Higgins, T. B., & Coggins, P. (2007). The increasing presence of undergraduate research in two-year colleges. *CUR Quarterly, 28*(2), 24-28.

Brownell, J. E., & Swaner, L. E. (2010). *Five high-impact practices: Research on learning outcomes, completion, and quality*. Association of American Colleges and Universities.

Butz, W. P., Kelly, T. K., Adamson, D. M., Bloom, G. A., Fossum, D., & Gross, M. E. (2004). *Will the scientific and technology workforce meet the requirements of the federal government?* RAND Corporation. https://www.rand.org/content/dam/rand/pubs/monographs/2004/RAND_MG118.pdf

Cejda, B. D., & Hensel, N. (Ed.). (2009). *Undergraduate research at community colleges*. Council on Undergraduate Research.

Cole, S. R., & Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology, 168*(6), 656-664. https://doi.org/10.1093/aje/kwn164

Council on Undergraduate Research (CUR). (2021). Undergraduate research at community colleges. *Scholarship and Practice of Undergraduate Research, 4*(3). The author.

Eagan, M. K., Jr., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013), Making a difference in science education: The impact of undergraduate research programs. *American Educational Research Journal, 50*(4), 683–713. https://dx.doi.org/10.3102%2F0002831213482038

Finley, A., & McNair, T. (2013). *Assessing underserved students engagement in high-impact practices*. Association of American Colleges and Universities.

Foley, C., & Leonhardt, N. (2017). Partnerships and collaborations impacting student research programs at community colleges. *Council on Undergraduate Research Quarterly, 37*(4), 4-11.

Genet, K. S. (2021). The CURE for introductory, large enrollment, and online courses. *Scholarship and Practice of Undergraduate Research, 4*(3), 13-21. doi: 10.18833/spur/4/3/14

Gentile, J., Brenner, K., & Stephens, A. (2017). *Undergraduate research experiences for STEM students: Successes, challenges, and opportunities*. National Academies Press.

Guo, S., & Fraser, M. W. (2015). *Advanced quantitative techniques in the social sciences: Vol. 11. Propensity score analysis: Statistical methods and applications* (2nd ed). SAGE Publications, Inc.

Hathaway, R., Nagda, B., & Gregerman, S. (2002). The relationship of undergraduate research participation to graduate and professional education pursuit: An empirical study. *Journal of College Student Development, 43*(5), 614-631.

Hensel, N. H. (2021). *Undergraduate research at community colleges: Equity, discovery, and innovation*. Stylus Publishing, LLC.

Hensel, N. H., & Cejda, B. D. (Eds.). (2014). *Tapping the potential of all: Undergraduate research at community colleges*. Council on Undergraduate Research.

Hewlett, J. A. (2016). Undergraduate research at the community college: Barriers and opportunities. *The Power and Promise of Early Research* (Chapter 8). ACS Symposium Series, 231, 137-151.

Hewlett, J. A. (2018). Broadening participation in undergraduate research experiences (UREs): The expanding role of the community college. *CBE–Life Sciences Education, 17*(3), es9. https://doi.org/10.1187/cbe.17-11-0238

Hewlett, J. (2021). Introduction: Undergraduate research at community colleges. *Scholarship and Practice of Undergraduate Research, 4*(3), 3-4. Council on Undergraduate Research. doi: 10.18833/spur/4/3/15

Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica, 71*(4), 1161–89.

Hirst, R., Bolduc, G., Liotta, L., & Packard, B. W. L. (2014). Cultivating the STEM transfer pathway and capacity for research: A partnership between a community college and a 4-year college. *Journal of College Science Teaching, 43*(4), 12-17.

Hunter, A-B., Laursen, S. L., & Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. *Science Education, 91*(1), 36-74. https://doi.org/10.1002/sce.20173

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics, 86*(1), 4-29.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments of the econometrics of program evaluation. *Journal of Economic Literature, 47*(1), 5-86.

Ishiyama J. (2001). Undergraduate research and the success of first generation, low income college students. *Council on Undergraduate Research Quarterly, 22*, 36-41.

Joffe, M. M., Ten Have, T. R., Feldman, H. I., & Kimmel, S. E. (2004). Model selection, confounder control, and marginal structural models: Review and new applications. *The American Statistician, 58*(4), 272–279. https://doi.org/10.1198/000313004X5824

Johnson, S. R., & Stage, F. K. (2018). Academic engagement and student success: Do high-impact practices mean higher graduation rates? *The Journal of Higher Education, 89*(5), 753-781. https://doi.org/10.1080/00221546.2018.1441107

Jones, M. T., Barlow, A., & Villarejo, M. (2010). Importance of undergraduate research for minority persistence and achievement in biology. *The Journal of Higher Education, 81*(1), 82–115. https://doi.org/10.1080/00221546.2010.11778971

Jonides, J., von Hippel, W., Lerner, J., & Nagda, B. (1992). *Evaluation of minority retention programs: The undergraduate research opportunities program at the University of Michigan*. Paper presented at the meeting of the American Psychological Association, Washington, DC.

Kilgo, C. A., Ezell Sheets, J. K., & Pascarella, E. T. (2015). The link between high-impact practices and student learning: Some longitudinal evidence. *Higher Education, 69*(4), 509–525. https://doi.org/10.1007/s10734-014-9788-z

Kuh, G. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter*. Association of American Colleges and Universities.

Laursen, S., Hunter, A. B., Seymour, E., Thiry, H., & Melton, G. (2010). *Undergraduate research in the sciences: Engaging students in real science.* Jossey-Bass.

Leggett-Robinson, P. M., Mooring, S. R., Villa, B. C. (2015). Two-year community: A 3+8 model of undergraduate research for community college STEM majors. *Journal of College Science Teaching, 44*(4), 12-18.

Linn, M. C., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: Impacts and opportunities. *Science, 347*(6222), 1261757. https://doi.org/10.1126/science.1261757

Lopatto, D. (2004). Survey of Undergraduate Research Experiences (SURE): First findings. *Cell Biology Education, 3*(4), 270-277. https://doi.org/10.1187/cbe.04-07-0045

Lopatto, D. (2006). Undergraduate research as a catalyst for liberal learning. *Peer Review*, 8(1), 22–25.

Lopatto, D. (2009). *Science in solution: The impact of undergraduate research on student learning*. The Research Corporation for Science Advancement.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted-regression for evaluating causal effects in observational studies. *Psychological Methods, 9*(4), 403-425. https://doi.org/10.1037/1082-989x.9.4.403

Nagda, B. A., Gregerman, S. R., Jonides, J., von Hippel, W., & Lerner, J. S. (1998). Undergraduate student-faculty research partnerships affect student retention. *Review of Higher Education, 22*(1), 55-72. DOI: 10.1353/rhe.1998.0016

Nerio, R., Webber, A., MacLachlan, E., Lopatto, D., & Caplan, A. J. (2019). One-year research experience for associate's degree students impacts graduation, STEM retention, and transfer patterns. *CBE—Life Sciences Education, 18*(2), 1-9. https://doi.org/10.1187/cbe.19-02-0042

Nolan, J. R., McConville, K. S., Addona, V., Tintle, N. L., & Pearl, D. K. (2020). Mentoring undergraduate research in statistics: Reaping the benefits and overcoming the barriers. *Journal of Statistics Education, 28*(2), 140-153. https://doi.org/10.1080/10691898.2020.1756542

O'Donnell, K., Botelho, J., Brown, J., González, G. M., & Head, W. (2015). Undergraduate research and its impact on student success for underrepresented students. *New Directions for Higher Education, 169*, 27-38. https://doi.org/10.1002/he.20120

Olsen, S. (Ed.). (2014). *The importance of engineering talent to the prosperity and security of the nation: Summary of a forum*. National Academies Press. https://doi.org/10.17226/18626

Pascarella, E. G., & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (Vol. 2). Jossey-Bass.

Patton, M., & Hause, E. (2020). *Community college undergraduate research experience summit proceedings report*. American Association of Community Colleges. http://www.aacc.nche.edu/URESummit

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in STEM*. Executive Office of the President.

Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association, 89*(427), 846–866.

Rodenbusch, S. E., Hernandez, P. R., Simmons, S. L., & Dolan, E. L. (2016). Early engagement in course-based research increases graduation rates and completion of science, engineering, and mathematics degrees. *CBE—Life Sciences Education, 15*(2), 1-10. https://doi.org/10.1187/cbe.16-03-0117

Rodríguez, A. L., Betancourt, T., Collins, C. H., Hinojosa, O., & Corona, C. (2018). Undergraduate research experiences: mentoring, awareness, and perceptions—a case study at a Hispanic-serving institution. *International Journal of STEM Education, 5*(9). https://doi.org/10.1186/s40594-018-0105-8

Russell, S. H. (2005). *Evaluation of NSF support for undergraduate research opportunities: 2003 NSF-Program Participant Survey: Final report*. SRI International.

Russell, S. H., Hancock, M. P., & McCullough J. (2007). Benefits of undergraduate research experiences. *Science*, *316*(5824), 548-549. DOI: 10.1126/science.1140384

Sato, T., & Matsuyama, Y. (2003). Marginal structural models as a tool for standardization. *Epidemiology*, *14*(6), 680-686. https://doi.org/10.1097/01.ede.0000081989.82616.7d

Schonlau, M. (2005). Boosted regression (boosting): An introductory tutorial and a Stata plugin. *The Stata Journal*, *5*(3), 330-354.

Schuster, M. (2017). Undergraduate research at two-year community colleges. *Journal of Political Science Education, 14*(2), 276-280. https://doi.org/10.1080/15512169.2017.1411273

Seymour, E., Hunter, A-B., Laursen, S. L., & DeAntoni, T. (2004). Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. *Science Education, 88*(4), 493-534. https://doi.org/10.1002/sce.10131

Snyder, J., & Cudney, E. A. (2017). Retention models for STEM majors and alignment to community colleges: A review of the literature. *Journal of STEM Education, 18*(3), 48-57.

Snyder, T. D., Dillow, S. A., & Staff, I. o. E. S. (2012). *Digest of Education Statistics 2011*. U.S. Government Printing Office.

Stanford, J. S., Rocheleau, S. E., Smith, K. P. W., & Mohan, J. (2017). Early undergraduate research experiences lead to similar learning gains for STEM and non-STEM undergraduates. *Studies in Higher Education, 42*(1), 115–129. https://doi.org/10.1080/03075079.2015.1035248

Summers, M. E., & Hrabowski III, E A. (2006). Preparing minority scientists and engineers. *Science, 311*(5769), 1870-1871. DOI: 10.1126/science.1125257

Tennessee Board of Regents (TBR). (2021). *The effect of service learning participation on college outcomes: An empirical investigation*. The author.

Tennessee Board of Regents. (n.d.). *HIP taxonomy: Undergraduate research and creative activities*. https://www.tbr.edu/student-success/hip-taxonomy-undergraduate-research-creative-activities

Thiry, H., & Laursen, S. L. (2011). The role of student-advisor interactions in apprenticing undergraduate researchers into a scientific community of practice. *Journal of Science Education and Technology, 20*(6), 771–784. https://doi.org/10.1007/s10956-010-9271-2

Thoemmes. F., & Ong, A. D. (2015). A primer on inverse probability of treatment weighting and marginal structural models. *Emerging Adulthood, 4*(1), 40-59. https://doi.org/10.1177/2167696815621645

Verity, P., Gilligan, M., Frischer, M., Booth, M., Richardson, J., & Franklin, C. (2002). Improving undergraduate research experiences: Lessons from a historically black university's unusual collaboration. *AAHE Bulletin*, *54*, 3-6.

Wolkow, T. D., Durrenberger, L. T., Maynard, M. A., Harrall, K. K., & Hines, L. M. (2014). A comprehensive faculty, staff, and student training program enhances student perceptions of a course-based research experience at a two-year institution. *CBE-Life Sciences Education, 13*(4), 724–737. https://doi.org/10.1187/cbe.14-03-0056

Zydney, A., Bennett, J., Shahid, A., & Bauer, K. (2002). Impact of undergraduate research on experience in engineering. *Journal of Engineering Education*, *91*(2), 151-157. https://doi.org/10.1002/j.2168-9830.2002.tb00687.x